Effective results for Diophantine problems over finitely generated domains

DSc dissertation

Attila Bérczes

University of Debrecen

2016

dc_1079_15

Acknowledgements

First of all I would like to thank my parents for their care, and especially my father, who directed my interest to mathematics already in my childhood. I would like to express my deep gratitude to my former PhD supervisor, Professor Kálmán Győry, without whose encouragement this dissertation would not be written, and to my co-authors, colleagues and teachers who supported and helped my work during many years. Finally, I would like to thank my wife and son for their love and for tolerating the long working hours I spent with research.

Köszönetnyilvánítás

Először is szeretném megköszönni szüleimnek a gondoskodást, és különösen Édesapámnak, hogy már gyermekkoromban felkeltette a matematika iránti érdeklődésemet. Ezúton fejezem ki mély hálámat korábbi PhD témavezetőmnek, Győry Kálmán Professzor Úrnak, akinek a bátorítása nélkül ez a diszszertáció nem készült volna el. Köszönöm a társszerzőimnek, kollégáimnak és tanáraimnak az évek során a munkámhoz nyújtott támogatását és segítséget. Végül, de nem utolsó sorban köszönöm feleségemnek és kisfiamnak a szeretetüket, és azt a türelmet, amivel folyamatosan elviselték a hosszú időt, amit kutatómunkával töltöttem.

dc_1079_15

Contents

Part I

Main Results

Chapter 1

Introduction

The three main problems in the theory of Diophantine equations are the decision of the solvability, studying the number of solutions and finding all solutions of the equation in question. In 1970 Matiasevich gave a negative answer to the famous 10th problem of Hilbert, namely, he proved that there does not exist an algorithm to decide the solvability of an arbitrary Diophantine equation. Thus those results which answer the three main problems for a wide class of Diophantine equations are of great importance.

Of particular importance are the so-called effective results, i.e. those results which not only state the finiteness of the number of solutions of a wide class of Diophantine equations, but also provide an algorithm for finding all solutions. In this respect the results of A. Baker (see [2], [3], [4], [5]) awarded by Fields Medal meant a great breakthrough. In this result Baker gave a non-trivial lower bound for linear forms in logarithms, which enabled him and others to prove effective results for various classes of Diophantine equations. These effective results are mainly of theoretical importance, generally they provide explicit upper bounds for the solutions in terms of the occurring parameters (e.g. degree and coefficients). These bounds are generally too large to directly enable us to solve the equations in question. However, in many cases the methods developed during the proof of these results combined with other methods, like the Lenstra-Lenstra-Lovász procedure, lead to practical algorithms, which make possible, using also computers, to completely solve such equations when the occurring parameters are of moderate sizes.

Classical Diophantine problems have been originally considered over the ring of rational integers. The results obtained over $\mathbb{Z}$ have had many applications. However, it turned out that even the study of rational integer solutions may need arguments over larger rings, like the ring of integers of an algebraic number field, or even more generally, a ring of S-integers of an algebraic number field. Thus it became important the investigation of solutions to

Diophantine problems coming from such rings. Such results lead to further applications among others in algebraic number theory. Finally, the development of the Diophantine geometry made necessary the extension of these results to equations considered over arbitrary finitely generated domains over $\mathbb{Z}$. Such domains may also contain transcendental elements over $\mathbb{Q}$. It is important to mention, that over domains which are not finitely generated such finiteness results are generally not true anymore.

Understandably, the first results obtained for a class of equations have been ineffective results, which did not provide any procedure (not even a theoretical one) to find all solutions of the equation in question. Later effective versions of these theorems have been established, however, in many cases in much less general contexts.

From the point of view of applications, among the most important classes of Diophantine equations are the *unit equations* and their *generalizations*, the *Thue-equation*, the *hyper- and superelliptic equation*, and the *Schinzel-Tijdeman equation*. In my present dissertation I present and prove my effective results concerning the above-mentioned classes of Diophantine equations considered over arbitrary finitely generated domains over $\mathbb{Z}$. All of the above types of equations have an extensive literature. Below I will recall the most important earlier achievements, then I summarize my results in a simplified form, indicating their place in the literature. The extensive presentation of my results will be done in Chapters 2 and 3, meanwhile the proofs are contained in Part II of the dissertation.

In the sequel let A be a finitely generated domain containing $\mathbb{Z}$ and let K denote the quotient field of A. Examples for such domains are $\mathbb{Z}$, more generally the ring of integers or S-integers of an algebraic number field, or even more generally a domain $\mathbb{Z}[z_1, \ldots, z_r]$, where $z_1, \ldots, z_r$ are either algebraic or transcendental elements over $\mathbb{Q}$, and also the polynomial rings $\mathbb{Z}[X_1, \ldots, X_r]$. It is a well-known fact, that the unit group A^* (i.e. the group of invertible elements) of such a domain A is a finitely generated group.

Following the terminology used by Lang, an equation of the form

$$ax + by = 1 \quad \text{in} \qquad x, y \in A^*, \tag{1.1}$$

where $a, b \in A \setminus \{0\}$, is called a *unit equation*. In the case when A is the ring of integers of an algebraic number field Siegel (1921) implicitly proved the finiteness of the number of solutions of equation (1.1). Mahler (1933) proved the finiteness for rings of the form $A = \mathbb{Z}[(p_1 \ldots p_s)^{-1}]$, where $p_1, \ldots, p_s$ are distinct primes. From results of Parry (1950) a common generalization over algebraic number fields of the theorems of Siegel and Mahler follows. Finally Lang (1960) proved the finiteness of the number of solutions in its full generality, over arbitrary finitely generated domains A.

Lang extended his result concerning equation (1.1) also to equations of the form

$$F(x,y)=0 \quad \text{in} \qquad x,y\in\Gamma, \tag{1.2}$$

where Γ is a finitely generated subgroup of K^*, and $F\in A[X,Y]$ is a polynomial which is not divisible by any polynomial of the shape

$$X^mY^n-\alpha \qquad \text{or} \qquad X^m-\alpha Y^n \tag{1.3}$$

with non-negative integers m,n not both zero, and $\alpha\in\Gamma$. It is easily seen that the conditions under (1.3) are in fact necessary. Lang [46], [47] (see also [48]) conjectured that his finiteness statement remains true in the more general case, when in (1.2) we take $\overline{\Gamma}$ instead of Γ, where

$$\overline{\Gamma}:=\left\{x\in\overline{K}^*\,:\,\exists m\in\mathbb{Z}_{>0}\ \text{ with }\ x^m\in\Gamma\right\}$$

is the division group of Γ. The conjecture of Lang has been proved by Liardet [51], [52].

The above-mentioned results are all ineffective, their proofs depend on the deep, however ineffective Thue-Siegel-Roth method.

The first effective results for the equation (1.1) over the ring of integers of algebraic number fields is due to Győry (1972, 1974), and over the ring of S-integers of an algebraic number field again due to Győry (1979). Using Baker's method he gave explicite upper bound for the heights of the solutions of (1.1). This bound has been improved by several authors. Later Mason [55] proved effective analogues of the effective result concerning equation (1.1) over function fields.

In special case, over algebraic number fields, Bombieri and Gubler [18] gave an effective version of Lang's theorem [44] on equation (1.2).

Győry (1983, 1984) developed a method to prove effective results for Diophantine equations considered over a wide class of finitely generated domains which may contain also transcendental elements. For unit equations and some other important classes of equations Győry proved his results over a larger domain B instead of A, using suitable effective specializations for the domain B and reducing his proof to the number field and function field case. However, with his method it was not possible to select the solutions belonging to A from those belonging to B. Recently Evertse and Győry (2013) combined the method of Győry with newer results of Aschenbrenner [1] and they proved an effective finiteness result for the unit equation (1.1) in full generality. They showed that equation (1.1) has finitely many and effectively determinable solutions for every finitely generated domain A both in A^*, and in any finitely generated subgroup Γ of K^*.

In the sequel I summarize the main results of my dissertation, which are effective finiteness results concerning equations (1.1) and (1.2).

In a joint work with Evertse and Győry [8] we extended earlier effective results for equation (1.1) *over algebraic number fields* to the case when the unknowns x, y belong to the division group of Γ, and even more generally, to a set containing points which are "close" to the division group of Γ; see Theorems 2.1, 2.3, 2.5 and Corollary 2.4. To prove these results we gave explicite lower bound for differences $|\alpha - \xi|_v$, where $\alpha \neq 0$ is an algebraic number, ξ belongs to a finitely generated multiplicative group containing only algebraic elements and v is an arbitrary place of a field containing α and the finitely generated group. Together with Evertse, Győry and Pontreau [11] we extended the results concerning equation (1.1) to the equation (1.2); see Theorems 2.8, 2.9 and 2.10. We also gave explicite upper bound for the height and degree of the solutions of equation (1.2) considerably generalizing and making explicite the result of Bombieri and Gubler [18].

Recently in [7] and [6] using the method of Evertse and Győry (but for the shape of the bound) I proved effective generalizations of all previous effective results concerning the solutions from Γ and $\overline{\Gamma}$ of the equations (1.1) and (1.2) *over arbitrary finitely generated domains* containing $\mathbb{Z}$. This is the main result of my dissertation.

There are many results concerning Thue-equations, and hyper- and superelliptic equations. Here I will only recall the most important results connected to our new achievements.

Consider first the Thue-equation. Let again A be a domain which is finitely generated over $\mathbb{Z}$, let $F(X, Y) \in A[X, Y]$ be a binary form of degree $n \geq 3$ and non-zero discriminant, let $\delta \in A \setminus \{0\}$, and consider the equation

$$F(x, y) = \delta \quad \text{in} \quad x, y \in A. \tag{1.4}$$

In the classical $A = \mathbb{Z}$ case Thue (1909) proved that equation (1.4) has only finitely many solutions. Thus equations of the type (1.4) are called *Thue-equations*. The result of Thue has been generalized by Siegel (1921) for the case when A is the ring of integers of an algebraic number field. Mahler (1933) extended Thue's Theorem for the case of rings of the shape $A = \mathbb{Z}[(p_1 \dots p_s)^{-1}]$, where $p_1, \dots, p_s$ are distinct primes. Parry (1950) gave a common generalization of the results of Siegel and Mahler. Finally Lang (1960) extended all these results to the case of arbitrary finitely generated domains A containing $\mathbb{Z}$. The proofs of the above-mentioned results depend on the Thue-Siegel-Roth method, so all these results are ineffective.

In the case $A = \mathbb{Z}$ the first general effective result was proved by Baker [2] using his new method based on lower bound for linear forms in logarithms. Coates [29] extended the result of Baker to the case of base rings of the type $A = \mathbb{Z}[(p_1 \dots p_s)^{-1}]$, and later

Kotov and Sprindžuk [71] to the case when A is the ring of S-integers of an algebraic number field. Mason [55] proved an effective function field analogue of the above results. Győry [39] using his effective specialization method generalized the above results for a wide but special class of finitely generated domains, which contained both algebraic and transcendental elements. We mention that the theory of unit equations and that of Thue equations is in fact equivalent; see e.g. Evertse and Győry [33].

Together with Evertse and Győry [10], using the already mentioned Evertse-Győry method [32] we obtained effective finiteness results for equation (1.4) in full generality, i.e. over arbitrary finitely generated domains A which contain $\mathbb{Z}$. This is Theorem 3.1 of the present dissertation.

Now consider the equation

$$F(x) = \delta y^m \qquad \text{in} \qquad x, y \in A, \tag{1.5}$$

where $F(X) \in A[X]$ is a polynomial of degree $n \geq 2$ with non-zero discriminant, $m \geq 2$ integer, and $\delta \in A \setminus \{0\}$. The equation (1.5) in the case $n \geq 3$, $m = 2$ is called *hyperelliptic equation*, and in the case $n \geq 2$, $m \geq 3$ it is called *superelliptic equation*. In the hyperelliptic case for $A = \mathbb{Z}$ Siegel (1926) obtained the first finiteness result. LeVeque (1964) gave a finiteness criterion for the equation (1.5) in the case when A is the ring of integers of an algebraic number field. Lang (1960) proved the finiteness of the number of solutions of (1.5) in full generality, i.e. over arbitrary finitely generated domains containing $\mathbb{Z}$.

For equation (1.5) in the case $A = \mathbb{Z}$ the first general effective result was also proved by Baker [3] using his effective method. Over $\mathbb{Z}$ Schinzel and Tijdeman [65] considered for the first case the equation (1.5) in the more general situation when also the exponent m is unknown, and they gave an effective upper bound for m. Thus in the case of unknown m the equation (1.5) is also called the *Schinzel-Tijdeman equation*. Trelina [74] and Brindza [21] gave independently effective upper bound for the height of the solutions of equation (1.5) over the ring of S-integers of an algebraic number field. Mason [55] proved an effective function field analogue of the above results. Finally, Brindza [22] and Végső [75] using the method developed by Győry [39], [40] proved effective finiteness results for equation (1.5) and the Schinzel-Tijdeman equation, respectively, over the special type of finitely generated domains considered also by Győry.

Together with Evertse and Győry [10] we obtained effective finiteness results for equation (1.5) and the Schinzel-Tijdeman equation in full generality, i.e. over arbitrary finitely generated domains A which contain $\mathbb{Z}$; see Theorems 3.3 and 3.4 of the present dissertation. Our proof is based again on the effective method developed by Evertse and Győry [32].

Here we mention that Evertse and Győry in their book [31] obtained effective finiteness results also for discriminant equations over arbitrary finitely generated domains A which contain $\mathbb{Z}$.

Altogether, one may say that the results of my dissertation are final, they conclude the effective investigation of equations (1.2), (1.4) and (1.5). These results are proved in quantitative form, i.e. they give effective upper bounds for the size of the solutions. Of course it remains an important task to decrease the bounds obtained for the solutions, and to develop practical, efficient algorithms which can be used for the complete solution of given equations in the case of parameters of moderate size.

Chapter 2

Results in the algebraic case

Among the finitely generated domains are of special importance those which contain only algebraic elements over $\mathbb{Q}$. In this chapter I summarize my effective results concerning the algebraic case of important Diophantine problems.

Choose an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. Recall that the group of $\overline{\mathbb{Q}}$-rational points of the N-dimensional torus is

$$\mathbb{G}_m^N(\overline{\mathbb{Q}}) = (\overline{\mathbb{Q}}^*)^N = \{\mathbf{x} = (x_1, \dots, x_N) : x_i \in \overline{\mathbb{Q}}^* \text{ for } i = 1, \dots, N\}$$

with coordinatewise multiplication, i.e., if $\mathbf{x} = (x_1, \dots, x_N)$, $\mathbf{y} = (y_1, \dots, y_N)$ then $\mathbf{xy} = (x_1y_1, \dots, x_Ny_N)$. Denote by $h(x)$ the absolute logarithmic Weil height of $x \in \overline{\mathbb{Q}}$. Define the height and degree of $\mathbf{x} = (x_1, \dots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ by $h(\mathbf{x}) := \sum_{i=1}^N h(x_i)$, and $[\mathbb{Q}(x_1, \dots, x_N) : \mathbb{Q}]$, respectively.

Let $\mathcal{X}$ be an algebraic subvariety of $(\overline{\mathbb{Q}}^*)^N$ (i.e., the set of common zeros in $(\overline{\mathbb{Q}}^*)^N$ of a set of polynomials in $\overline{\mathbb{Q}}[X_1, \dots, X_N]$), and Γ a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$.

We want to study the intersection of $\mathcal{X}$ with any of the sets

$$\begin{aligned}
&\overline{\Gamma} := \Big\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \exists m \in \mathbb{Z}_{>0} \text{ with } \mathbf{x}^m \in \Gamma\Big\} \quad \text{(the division group of } \Gamma\text{)},\\
&\overline{\Gamma}_\varepsilon := \Big\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \exists \mathbf{y}, \mathbf{z} \in (\overline{\mathbb{Q}}^*)^N \text{ with } \mathbf{x} = \mathbf{yz},\ \mathbf{y} \in \overline{\Gamma},\ h(\mathbf{z}) < \varepsilon\Big\},\\
&C(\overline{\Gamma}, \varepsilon) := \Big\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \exists \mathbf{y}, \mathbf{z} \in (\overline{\mathbb{Q}}^*)^N\\
&\qquad\qquad \text{with } \mathbf{x} = \mathbf{yz},\ \mathbf{y} \in \overline{\Gamma},\ h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y}))\Big\},
\end{aligned} \tag{2.1}$$

where $\varepsilon > 0$. We mention, that in contrast to $\overline{\Gamma}$ which is a group, the sets $\overline{\Gamma}_\varepsilon$ and $C(\overline{\Gamma}, \varepsilon)$ do not form any algebraic structure with respect to the operations inherited from $(\overline{\mathbb{Q}})^N$.

Recall that by an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$ we mean an algebraic subvariety that is

a subgroup of $(\overline{\mathbb{Q}}^*)^N$, and by a translate of an algebraic subgroup a coset $\mathbf{x}\mathcal{H} = \{\mathbf{x} \cdot \mathbf{y} : \mathbf{y} \in \mathcal{H}\}$, where $\mathcal{H}$ is an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$ and $\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N$.

Concerning the above-mentioned intersections several **ineffective** results have been published. It follows from work of Poonen [60] that there is $\varepsilon > 0$ depending only on N and the degree of $\mathcal{X}$, such that $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$ is contained in a finite union of translates

$$\mathbf{x}_1\mathcal{H}_1 \cup \cdots \cup \mathbf{x}_T\mathcal{H}_T \tag{2.2}$$

where $\mathbf{x}_i \in \overline{\Gamma}_\varepsilon$, $\mathcal{H}_i$ is an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$ and $\mathbf{x}_i\mathcal{H}_i \subset \mathcal{X}$ for $i = 1, \ldots, T$. This encompasses earlier work of Liardet [52] and Laurent [49] (who considered $\mathcal{X} \cap \overline{\Gamma}$) and Zhang [79] (who considered $\mathcal{X} \cap \{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : h(\mathbf{x}) < \varepsilon\}$).

Bombieri and Zannier [19] and Schmidt [69] proved precise quantitative versions for Zhang's result with an explicit positive value for ε and an explicit upper bound for the number T of translates, both depending only on N and the degree of $\mathcal{X}$ and their result was further improved by various authors. Later, Rémond [61] proved a quantitative version of Poonen's result with an explicit positive value for ε depending on N and the degree of $\mathcal{X}$ and an explicit upper bound for T depending only on N, the degree of $\mathcal{X}$ and the rank of Γ.

Define $\mathcal{X}^{\mathrm{exc}}$ to be the set of $\mathbf{x} \in \mathcal{X}$ with the property that there exists an algebraic subgroup $\mathcal{H}$ of $(\overline{\mathbb{Q}}^*)^N$ of dimension > 0 such that $\mathbf{x}\mathcal{H} \subset \mathcal{X}$, and let $\mathcal{X}^0 := \mathcal{X} \setminus \mathcal{X}^{\mathrm{exc}}$. Evertse stated in the survey paper [30] that there exists $\varepsilon > 0$ depending on N, $\mathcal{X}$ and Γ such that $\mathcal{X}^0 \cap C(\overline{\Gamma}, \varepsilon)$ is finite. This was proved in a more general form by Rémond [61]. In the case that $\mathcal{X}$ is a curve, Rémond gave, for some explicit value of ε depending on N, the rank of Γ and the height and degree of $\mathcal{X}$, an explicit upper bound for the cardinality of $\mathcal{X}^0 \cap C(\overline{\Gamma}, \varepsilon)$; his result was improved by Pontreau [58] for curves in $(\overline{\mathbb{Q}}^*)^2$. For higher dimensional varieties, such a quantitative version has as yet not been established.

In the present chapter we derive, for certain special classes of varieties $\mathcal{X}$, **effective** versions of the results mentioned above. As for the intersection $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$, this means that we give an explicit value for ε and effectively computable upper bounds for the heights and degrees of the points $\mathbf{x}_1, \ldots, \mathbf{x}_T$ in (2.2). As for $\mathcal{X}^0 \cap C(\overline{\Gamma}, \varepsilon)$, this means that we give an explicit value for ε and effectively computable upper bounds for the heights and degrees of the points in this intersection. We mention that to obtain fully effective results it is necessary to give effective upper bounds for the degrees as well since the points we are considering do not have their coordinates in a prescribed algebraic number field.

The classes of varieties we consider are such that they allow an application of non-trivial lower estimates of logarithmic forms, i.e. Baker's method. Three cases are worked out in detail. Firstly, in the case $N = 2$, if the variety is defined by a single linear polynomial,

then in principle we are dealing with the solutions of a generalized unit equation in two unknowns. These results are presented in the second section of the present chapter. Secondly, again in the special case $N = 2$, we consider curves $\mathcal{C} : f(x, y) = 0$ in $(\overline{\mathbb{Q}}^*)^2$ where $f \in \overline{\mathbb{Q}}[X, Y]$ is not a binomial. These results are contained in the third section. Finally, we consider varieties in $(\overline{\mathbb{Q}}^*)^N$ given by equations $f_1(\mathbf{x}) = 0, \ldots, f_m(\mathbf{x}) = 0$ where each polynomial f_i is a binomial or trinomial. These results form the fourth section of the present chapter, and their proofs depend profoundly on our results on the equation $ax + by = 1$ presented in the second section.

2.1 Notations

Let K be an algebraic number field of degree d. Denote by $\mathcal{O}_K$ its ring of integers and by M_K its set of places. For $v \in M_K$, we define an absolute value $|\cdot|_v$ as follows. If v is infinite and corresponds to $\sigma : K \to \mathbb{C}$, then we put $|x|_v = |\sigma(x)|^{d_v/d}$ for $x \in K$, where $d_v = 1$ or 2 according as $\sigma(K)$ is contained in $\mathbb{R}$ or not; if v is a finite place corresponding to a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, then we put $|x|_v = N(\mathfrak{p})^{-\operatorname{ord}_\mathfrak{p} x/d}$ for $x \in K \setminus \{0\}$, and $|0|_v = 0$. Here $N(\mathfrak{p})$ denotes the norm of $\mathfrak{p}$, and $\operatorname{ord}_\mathfrak{p} x$ the exponent of $\mathfrak{p}$ in the prime ideal factorization of the principal fractional ideal (x).

The absolute logarithmic height $h(x)$ of $x \in K$ is defined by

$$h(x) = \sum_{v \in M_K} \max(0, \log |x|_v). \tag{2.3}$$

More generally, if $x \in \overline{\mathbb{Q}}$ then choose an algebraic number field K such that $x \in K$ and define $h(x)$ by (2.3). This definition does not depend on the choice of K. Notice that $h(x) = 0$ if and only if $x \in \overline{\mathbb{Q}}^*_{\text{tors}}$, where $\overline{\mathbb{Q}}^*_{\text{tors}}$ is the group of roots of unity in $\overline{\mathbb{Q}}^*$.

Let S denote a finite subset of M_K containing all infinite places. Then $x \in K$ is called an S-integer if $|x|_v \leq 1$ for all $v \in M_K \setminus S$. More precisely, we define the ring of S-integers and group of S-units by

$$\begin{aligned} \mathcal{O}_S &= \{x \in K : |x|_v \leq 1 \text{ for } v \in M_K \setminus S\}, \\ \mathcal{O}_S^* &= \{x \in K : |x|_v = 1 \text{ for } v \in M_K \setminus S\}, \end{aligned}$$

respectively. In fact $\mathcal{O}_S^*$ is the unit group of the ring $\mathcal{O}_S$.

For every $v \in M_K$ put

$$P(v) := 2 \text{ if } v \text{ is infinite}, \qquad P(v) := \#\mathcal{O}_K/\mathfrak{p}_v \text{ if } v \text{ is finite}, \tag{2.4}$$

where $\mathfrak{p}_v$ is the prime ideal of $\mathcal{O}_K$ which corresponds to the place v. Put further

$$\mathbf{P} := \max_{v \in S} P(v). \tag{2.5}$$

It follows from (2.3) and the product formula that

$$h(x) = \frac{1}{2} \sum_{v \in S} |\log |x|_v| \quad \text{if} \quad x \in \mathcal{O}_S^*. \tag{2.6}$$

We define the height of $\mathbf{x} = (x_1, \dots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ by

$$h(\mathbf{x}) := \sum_{i=1}^{n} h(x_i).$$

For $\xi \in \mathbb{Q}$ let $\mathbf{x}^\xi := (x_1^\xi, \dots, x_N^\xi)$. The point $\mathbf{x}^\xi$ is determined only up to multiplication by elements from $(\overline{\mathbb{Q}}^*_{\text{tors}})^N$, where $\overline{\mathbb{Q}}^*_{\text{tors}} = \{\boldsymbol{\rho} \in \overline{\mathbb{Q}}^* \ : \ \exists m \in \mathbb{Z}_{>0} \ \text{ with } \ \boldsymbol{\rho}^m = 1\}$. However, the height $h(\mathbf{x}^\xi)$ is well defined, and we have:

$$h(\mathbf{xy}) \le h(\mathbf{x}) + h(\mathbf{y}), \ h(\mathbf{x}^\xi) = |\xi| h(\mathbf{x}) \qquad \text{for every } \mathbf{x}, \mathbf{y} \in (\overline{\mathbb{Q}}^*)^N \text{ and } \xi \in \mathbb{Q}.$$

Further $h(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in (\overline{\mathbb{Q}}^*_{\text{tors}})^N$.

For a number field L and for $\mathbf{x} = (x_1, \dots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ we define the extension $L(\mathbf{x}) := L(x_1, \dots, x_N)$.

For a polynomial $f \in \overline{\mathbb{Q}}[X_1, \dots, X_N]$ let $\deg f$ denote its total degree and put $\deg_s f := \sum_{i=1}^N \deg_{X_i} f$, where $\deg_{X_i}$ is the degree of f in the variable X_i. Assume that the non-zero coefficients of f are $a_1, \dots, a_R$ and put $K := \mathbb{Q}(a_1, \dots, a_R)$. Then the height of f is defined by

$$h(f) := \sum_{v \in M_K} \log \max_{1 \le i \le R} |a_i|_v.$$

Let $\log^* x := \max(1, \log x)$ for every $x > 0$ and put $\log^* 0 := 1$.

2.2 Effective results for generalized unit equations

In this section we consider the case $N = 2$ in the important special case, when the variety $\mathcal{X}$ is defined by a single linear polynomial. In this case the intersection of $\mathcal{X}$ with the sets $\overline{\Gamma}, \overline{\Gamma}_\varepsilon, C(\overline{\Gamma}, \varepsilon)$ is the set of solutions in $\overline{\Gamma}, \overline{\Gamma}_\varepsilon, C(\overline{\Gamma}, \varepsilon)$, respectively, of the Diophantine equation induced by the polynomial which defines $\mathcal{X}$. Since now $\mathcal{X}$ is defined by a linear polynomial, thus in principle we are analyzing the solutions of a generalized S-unit equation coming from the set $\overline{\Gamma}, \overline{\Gamma}_\varepsilon, C(\overline{\Gamma}, \varepsilon)$, respectively.

In the literature there are various effective results on S-unit equations in two unknowns. In this Section we present effective results in a quantitative form for the more general equation

$$a_1x_1 + a_2x_2 = 1 \qquad \text{in } (x_1, x_2) \in \Gamma, \tag{2.7}$$

where $a_1, a_2 \in \overline{\mathbb{Q}}^*$ and Γ is an arbitrary finitely generated subgroup of rank > 0 of the multiplicative group $(\overline{\mathbb{Q}}^*)^2 = \overline{\mathbb{Q}}^* \times \overline{\mathbb{Q}}^*$ endowed with coordinatewise multiplication (see Theorems 2.1 and 2.2 in the present Section). Such more general results can be used to improve upon existing effective bounds on the solutions of discriminant equations and certain decomposable form equations.

In fact, in the present section we present even more general effective results for equations of the shape (2.7) with solutions (x_1, x_2) from a larger group, from the division group $\overline{\Gamma} = \{(x_1, x_2) \in (\overline{\mathbb{Q}}^*)^2 \mid \exists k \in \mathbb{Z}_{>0} : (x_1^k, x_2^k) \in \Gamma\}$, and even with solutions (x_1, x_2) 'very close' to $\overline{\Gamma}$. These results give an effective upper bound for both the height of a solution (x_1, x_2) and the degree of the field $\mathbb{Q}(x_1, x_2)$; see Theorems 2.3 and 2.5 and Corollary 2.4.

In the proofs of these Theorems we utilize Theorem 2.1 (on (2.7) with solutions from Γ), as well as a result of Beukers and Zagier [14], which asserts that (2.7) has at most two solutions $(x_1, x_2) \in (\overline{\mathbb{Q}}^*)^2$ with very small height.

The hard core of the proofs of our results mentioned above is a new effective lower bound for $|1 - \alpha\xi|_v$, where α is a fixed element from a given algebraic number field K, v is a place of K, and the unknown ξ is taken from a given finitely generated subgroup of K^* (see Theorem 2.6). This result is proved using linear forms in logarithms estimates. Our Theorem 2.6 has a consequence (cf. Theorem 2.7) which is of a similar flavour as earlier results by Bombieri [15], Bombieri and Cohen [16], [17], and Bugeaud [26] (see also Bombieri and Gubler [18, Chap. 5.4]) but it gives in many cases a better estimate. Consequently, Theorem 2.6 leads to an explicit upper bound for the heights of the solutions of (2.7) which is in many cases sharper than what is obtainable from the work of Bombieri et al.

We consider again the equation

$$a_1x_1 + a_2x_2 = 1 \quad \text{in } (x_1, x_2) \in \Gamma \tag{2.7}$$

where $a_1, a_2 \in \overline{\mathbb{Q}}^*$ and where Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^2$ of rank > 0. Let $\mathbf{w}_1 = (\xi_1, \eta_1), \ldots, \mathbf{w}_r = (\xi_r, \eta_r)$ be a system of generators of $\Gamma/\Gamma_{\text{tors}}$ which is not necessarily a basis. Notice that every element of Γ can be expressed as $\zeta\mathbf{w}_1^{x_1} \cdots \mathbf{w}_r^{x_r}$ where $x_1, \ldots, x_r \in \mathbb{Z}$ and $\zeta \in \Gamma_{\text{tors}}$, i.e., the coordinates of ζ are roots of unity.

Define $K := \mathbb{Q}(\Gamma)$, i.e. the field generated by Γ over $\mathbb{Q}$. We do not require that $a_1, a_2 \in K$. Let S be the smallest set of places of K containing all infinite places such that

$\mathbf{w}_1, \dots, \mathbf{w}_r \in (\mathcal{O}_S^*)^2$, where $\mathcal{O}_S^*$ denotes the group of S-units in K. Put

$$Q_\Gamma := h(\mathbf{w}_1) \cdots h(\mathbf{w}_r),$$
$$d := [K : \mathbb{Q}], \quad s := \#S, \quad \mathbf{P} := \max_{v \in S} \mathbf{P}(v).$$

Denote by t the maximum of the rank of the subgroup of $\overline{\mathbb{Q}}^*$ generated by $\xi_1, \dots, \xi_r$ and the rank of the subgroup generated by $\eta_1, \dots, \eta_r$. In view of $\operatorname{rank} \Gamma > 0$ we have $t > 0$. We define

$$\begin{aligned} c_1(r,d,t) &:= 3(16ed)^{3(t+2)} \left(d(\log 3d)^3\right)^{r-t} (t/e)^t, \\ A &:= 26 c_1(r,d,t) s \frac{\mathbf{P}}{\log \mathbf{P}} Q_\Gamma \max\{\log(c_1(r,d,t)s\mathbf{P}), \log^* Q_\Gamma\}, \\ H &:= \max(h(a_1), h(a_2), 1). \end{aligned} \tag{2.8}$$

Then our first result reads as follows:

Theorem 2.1 (Bérczes, Evertse and Győry [8]). *For every solution $(x_1, x_2) \in \Gamma$ of (2.7) we have*

$$h(x_1, x_2) < AH. \tag{2.9}$$

We shall deduce Theorem 2.1 from Theorem 2.2 below. Let G be a finitely generated multiplicative subgroup of $\overline{\mathbb{Q}}^*$ of rank $t > 0$, and $\xi_1, \dots, \xi_r$ a system of generators of G/G_{tors}. Let K be a number field containing G, and S a finite set of places of K containing the infinite places such that $G \subseteq \mathcal{O}_S^*$. We consider the equation

$$a_1 x_1 + a_2 x_2 = 1 \qquad \text{in } x_1 \in G,\ x_2 \in \mathcal{O}_S^*, \tag{2.10}$$

where $a_1, a_2 \in \overline{\mathbb{Q}}^*$. Let $d := [K : \mathbb{Q}]$. Let s be the cardinality of S, $\mathbf{P} := \max_{v \in S} P(v)$ and put

$$Q_G := h(\xi_1) \cdots h(\xi_r).$$

Theorem 2.2 (Bérczes, Evertse and Győry [8]). *Under the above assumptions and notation, every solution of (2.10) satisfies*

$$h(x_1) < c_1(r,d,t) s \frac{\mathbf{P}}{\log \mathbf{P}} Q_G H \log^* \left(\frac{\mathbf{P} h(x_1)}{H} \right) \tag{2.11}$$

and

$$\begin{aligned} \max(h(x_1), h(x_2)) < 6.5 c_1(r,d,t) s \frac{\mathbf{P}}{\log \mathbf{P}} Q_G H \cdot \\ \cdot \max\{\log(c_1(r,d,t)s\mathbf{P}), \log^* Q_G\}, \end{aligned} \tag{2.12}$$

where as before $H := \max(h(a_1), h(a_2), 1)$ and $c_1(r,d,t)$ is the constant defined in (2.8).

If in particular $r = t$ and $\{\xi_1, \dots, \xi_t\}$ is a basis of G/G_{tors}, then, in (2.11) and (2.12) we can replace $c_1(r,d,t)$ by $c_1(d,t) = 73(16ed)^{3t+5}$.

As a special case of our result, it follows an effective upper bound for heights of the solutions of S-unit equations. As we mentioned it previously, the first such effective bound was proved by Győry [38], and it was improved by several mathematicians. The presently known best bounds are due to Bugeaud and Győry [27], Bugeaud [26], and Győry and Yu [41]. From the proof of our result it is possible to deduce an upper bound, which is comparable with these best known results.

We now consider equations such as (2.7) but with solutions (x_1, x_2) from a larger set. We keep the notation introduced before Theorem 2.1.

The division group of Γ is given by

$$\overline{\Gamma} := \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^2 \mid \exists k \in \mathbb{Z}_{>0} \text{ with } \mathbf{x}^k \in \Gamma \right\}.$$

For any $\varepsilon > 0$ define the "cylinder" and "truncated cone" around $\overline{\Gamma}$ by

$$\overline{\Gamma}_\varepsilon := \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^2 \mid \exists \mathbf{y}, \mathbf{z} \text{ with } \mathbf{x} = \mathbf{yz},\ \mathbf{y} \in \overline{\Gamma},\ \mathbf{z} \in (\overline{\mathbb{Q}}^*)^2,\ \ h(\mathbf{z}) < \varepsilon \right\} \tag{2.13}$$

and

$$\begin{aligned} C(\overline{\Gamma}, \varepsilon) := \Big\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^2 \mid \exists \mathbf{y}, \mathbf{z} \text{ with } \mathbf{x} = \mathbf{yz},\ \mathbf{y} \in \overline{\Gamma}, \\ \mathbf{z} \in (\overline{\mathbb{Q}}^*)^2,\ h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y})) \Big\}, \end{aligned} \tag{2.14}$$

respectively. The set $\overline{\Gamma}_\varepsilon$ was introduced by Poonen [60] and the set $C(\overline{\Gamma}, \varepsilon)$ by Evertse [30] (both in a much more general context).

We emphasize that points from $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$ do not have their coordinates in a prescribed number field. So for effective results on Diophantine equations with solutions from $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$, we need an effective upper bound not only for the height of each solution, but also for the degree of the field which it generates. We fix $a_1, a_2 \in \overline{\mathbb{Q}}^*$ and define

$$K := \mathbb{Q}(\Gamma), \quad K_0 := \mathbb{Q}(a_1, a_2, \Gamma).$$

The quantities $d, s, \mathbf{P}, H$ and Q_Γ will have the same meaning as in Theorem 2.1 and A will be the constant defined in (2.8). Further, we put

$$h_0 := \max\{h(\xi_1), \ldots, h(\xi_r), h(\eta_1), \ldots, h(\eta_r)\},$$

where $\mathbf{w}_i = (\xi_i, \eta_i)$ for $i = 1, \ldots, r$ is the chosen system of generators for $\Gamma / \Gamma_{\text{tors}}$.

Consider now the equation

$$a_1 x_1 + a_2 x_2 = 1 \qquad \text{in } (x_1, x_2) \in \overline{\Gamma}_\varepsilon. \tag{2.15}$$

Theorem 2.3 (Bérczes, Evertse and Győry [8]). *Suppose that (x_1, x_2) is a solution of (2.15) and that*

$$\varepsilon < 0.0225. \tag{2.16}$$

Then we have

$$h(x_1, x_2) \leq Ah(a_1, a_2) + 3rh_0A \tag{2.17}$$

and

$$[K_0(x_1, x_2) : K_0] \leq 2. \tag{2.18}$$

The following consequence is immediate:

Corollary 2.4 (Bérczes, Evertse and Győry [8]). *With the above notation and assumptions, let (x_1, x_2) be a solution of*

$$a_1x_1 + a_2x_2 = 1 \qquad \text{in } (x_1, x_2) \in \overline{\Gamma}. \tag{2.19}$$

Then $h(x_1, x_2) \leq Ah(a_1, a_2) + 3rh_0A$ and $[K_0(x_1, x_2) : K_0] \leq 2$.

Finally we consider the equation

$$a_1x_1 + a_2x_2 = 1 \qquad \text{in } (x_1, x_2) \in C(\overline{\Gamma}, \varepsilon). \tag{2.20}$$

Theorem 2.5 (Bérczes, Evertse and Győry [8]). *Suppose that (x_1, x_2) is a solution of (2.20) and that*

$$\varepsilon < \frac{0.09}{8Ah(a_1, a_2) + 20rh_0A}. \tag{2.21}$$

Then we have

$$h(x_1, x_2) \leq 3Ah(a_1, a_2) + 5rh_0A \tag{2.22}$$

and

$$[K_0(x_1, x_2) : K_0] \leq 2. \tag{2.23}$$

The paper [12] of Beukers and Schlickewei gives explicit upper bounds for the number of solutions of (2.19), while from the result of Evertse, Schlickewei and Schmidt [34] one can deduce explicit upper bounds for multivariate generalizations of (2.19), (2.15), (2.20). The sets $\overline{\Gamma}_\varepsilon$ and $C(\overline{\Gamma}, \varepsilon)$ have been defined in the much more general context of semi-abelian varieties (see [60], [63]). In [61], [62], Rémond proved quantitative analogues of the work of Evertse, Schlickewei and Schmidt [34] for subvarieties of abelian varieties and subvarieties of tori. We mention that the results of [12], [34], [60], [61], [62] and [63] are all ineffective, i.e. it does not even provide a theoretical algorithm to find the solutions. In contrast, our

above presented results are effective, i.e. they provide a theoretical algorithm to find all solutions of the equations in question, even though this algorithm is not practical.

The proof of the Theorems presented in this section is based on our two Diophantine approximation theorems presented below.

Let again K be an algebraic number field of degree d, M_K the set of places on K, and G a finitely generated multiplicative subgroup of K^* of rank $t > 0$. Further, let $\{\xi_1, \dots, \xi_r\}$ be a system of (not necessarily multiplicatively independent) generators of G such that $\xi_1, \dots, \xi_r$ are not roots of unity. Put again

$$Q_G := h(\xi_1) \cdots h(\xi_r).$$

Further, for any $v \in M_K$ let $P(v)$ be as in (2.4).

Theorem 2.2 and then subsequently Theorem 2.1 will be deduced from the following theorem.

Theorem 2.6 (Bérczes, Evertse and Győry [8]). *Let $\alpha \in K^*$ with $\max(h(\alpha), 1) \le H$ and let $v \in M_K$. Then for every $\xi \in G$ for which $\alpha\xi \neq 1$, we have*

$$\log |1 - \alpha\xi|_v > -c_2(r,d,t) \frac{P(v)}{\log P(v)} Q_G H \log^* \left(\frac{P(v)h(\xi)}{H} \right), \tag{2.24}$$

where

$$c_2(r,d,t) = (16ed)^{3(t+2)} \left(d(\log 3d)^3 \right)^{r-t} (t/e)^t.$$

In particular, if $r = t$ and $\{\xi_1, \dots, \xi_t\}$ is a basis of G/G_{tors}, then (2.24) holds with $c_2(d,t) = 36(16ed)^{3t+5}(\log^* d)^2$ instead of $c_2(r,d,t)$.

It should be observed that $c_2(d,t)$ does not contain a t^t factor.

The following theorem is in fact an immediate consequence of Theorem 2.6.

Theorem 2.7 (Bérczes, Evertse and Győry [8]). *Let $\alpha \in K^*$ with $\max(h(\alpha), 1) \le H$, let $v \in M_K$, and let $0 < \kappa \le 1$. Then for every $\xi \in G$ with $\alpha\xi \neq 1$ and*

$$\log |1 - \alpha\xi|_v < -\kappa h(\xi) \tag{2.25}$$

we have

$$h(\xi) < (c_2(r,d,t)/\kappa) \frac{P(v)}{\log P(v)} Q_G H \log^* \left(\frac{P(v)h(\xi)}{H} \right) \tag{2.26}$$

and

$$\begin{aligned} h(\xi) < 6.4(c_2(r,d,t)/\kappa) \frac{P(v)}{\log P(v)} Q_G H \cdot \\ \cdot \max \left(\log \left((c_2(r,d,t)/\kappa) P(v) \right), \log^* Q_G \right) \end{aligned} \tag{2.27}$$

with the constant $c_2(r,d,t)$ specified in Theorem 2.6.

In particular, if $r = t$ and $\{\xi_1, \dots, \xi_t\}$ is a basis of G/G_{tors}, then (2.26) and (2.27) hold with $c_2(d,t)$ instead of $c_2(r,d,t)$.

We note that when applying Theorem 2.7 to equation (2.10), inequality (2.26) yields better bounds in Theorem 2.2 than (2.27).

The main tool in the proofs of Theorems 2.6 and 2.7 is the theory of logarithmic forms, i.e. Baker's method, more precisely Theorem C in Section 4.1. Bombieri [15] and Bombieri and Cohen [16], [17] have developed another effective method in Diophantine approximation, based on an extended version of the Thue-Siegel principle, the Dyson lemma and some geometry of numbers. Bugeaud [26], following their approach and combining it with estimates for linear forms in two and three logarithms, obtained sharper results than Bombieri and Cohen. Our above result is comparable with that of Bugeaud, and in most of the parameters (except in some cases the parameters H and Q_G) it is sharper than the result of Bugeaud.

2.3 Generalized unit points on curves

In this section we consider curves given by $\mathcal{C} := \{(x,y) \in (\overline{\mathbb{Q}}^*)^2 \mid f(x,y) = 0\}$ where $f(X,Y)$ is a polynomial which is not a binomial. We generalize on one hand results of Bombieri and Gubler [18, p. 147, Theorem 5.4.5], on the other hand results of Bérczes, Evertse and Győry [8, Theorems 2.1, 2.3 and 2.5]. These latter results appear as Theorem 2.1, 2.3 and 2.5 in the second section of the present chapter. More precisely we give upper bounds for the height and degree of those points $\mathbf{x}$ which are contained in the intersection of $\mathcal{C}$ with one of the sets Γ, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$. Our result concerning $\mathcal{C} \cap \overline{\Gamma}$ is in fact the first effective version of the famous finiteness theorem of Liardet [51], [52] for division points on curves, however only in the special case when Γ contains only algebraic elements.

Let Γ be a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^2$. Further, let $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ and $C(\overline{\Gamma}, \varepsilon)$ be defined as in (2.1) in the special case $N = 2$. Choose a basis $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$ of Γ modulo Γ_{tors} and put

$$h_0 := \max\big(1, h(\mathbf{w}_1), \dots, h(\mathbf{w}_r)\big).$$

Denote by K the smallest number field such that $\Gamma \subset (K^*)^2$, and put $d := [K : \mathbb{Q}]$. Let S be the minimal finite set of places of K containing all the infinite places of K and having the property that $\Gamma \subset (\mathcal{O}_S^*)^2$ and denote by s the cardinality of S. Let $\mathbf{P}$ be the quantity defined in (2.5).

Let $f(X,Y) \in \overline{\mathbb{Q}}[X,Y]$ be an absolutely irreducible polynomial which is not of the shape $aX^mY^n - b$ or $aX^m - bY^n$ for some $a, b \in \overline{\mathbb{Q}}$, $m, n \in \mathbb{Z}_{\geq 0}$. Let L be the field extension of K generated by the coefficients of f. Put

$$\delta := \deg_s f, \quad H := \max(1, h(f)),$$
$$C_1 := \left(e^{13}\delta^7 d^3 r\right)^{r+3} s \cdot \frac{\mathbf{P}^{2\delta^2}}{\log \mathbf{P}} h_0^r \cdot \log^* \left(\max(\delta d s \mathbf{P}, \delta h_0)\right).$$

Let $\mathcal{C} \subset (\overline{\mathbb{Q}}^*)^2$ be the curve defined by $f(x,y) = 0$. By our assumptions on f, $\mathcal{C}$ is not a translate of a proper algebraic subgroup of $(\overline{\mathbb{Q}}^*)^2$.

Theorem 2.8 (Bérczes, Evertse, Győry and Pontreau [11]). *For every point* $\mathbf{x} = (x,y) \in \mathcal{C} \cap \Gamma$ *we have*

$$h(\mathbf{x}) = h(x) + h(y) \leq C_1 H.$$

Notice that in this bound there is no dependence on the field L other than what is implicit from H.

The following results are obtained by combining the above theorem with estimates for the number of points of small height on a curve in $(\overline{\mathbb{Q}}^*)^2$. The notation will be the same as above.

Theorem 2.9 (Bérczes, Evertse, Győry and Pontreau [11]). *Let*

$$\varepsilon := \left(2^{48}\delta(\log \delta)^5\right)^{-1}. \tag{2.28}$$

Then for every $\mathbf{x} \in \mathcal{C} \cap \overline{\Gamma}_\varepsilon$ *we have*

$$h(\mathbf{x}) \leq r h_0 \delta C_1 + C_1 H, \quad [L(\mathbf{x}) : L] \leq 2^{50}\delta(\log \delta)^6.$$

Theorem 2.10 (Bérczes, Evertse, Győry and Pontreau [11]). *Let*

$$\varepsilon := \left(2^{50}\delta(\log \delta)^5\right)^{-1} \cdot \left(r h_0 \delta C_1 + C_1 H\right)^{-1}. \tag{2.29}$$

Then for every $\mathbf{x} \in \mathcal{C} \cap C(\overline{\Gamma}, \varepsilon)$ *we have*

$$h(\mathbf{x}) \leq 2 r h_0 \delta C_1 + 2 C_1 H, \quad [L(\mathbf{x}) : L] \leq 2^{50}\delta(\log \delta)^6.$$

Remark. In the special case when f is linear, (i.e., $\mathcal{C}$ is a line), our above theorems have been proved in [8] with larger ε's and sharper upper bounds.

2.4 Generalized unit points on N-dimensional varieties

In this chapter we turn our attention to varieties of arbitrary dimension N.

Let Γ be a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$, where $N \geq 2$. Further, let $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ and $C(\overline{\Gamma}, \varepsilon)$ be defined as in (2.1). Choose a basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ of Γ modulo Γ_{tors} and put

$$h_0 := \max\big(1, h(\mathbf{w}_1), \ldots, h(\mathbf{w}_r)\big).$$

Denote by K the smallest number field such that $\Gamma \subset (K^*)^N$, and put $d := [K : \mathbb{Q}]$. Let S be the minimal finite set of places of K containing all the infinite places of K and having the property that $\Gamma \subset (\mathcal{O}_S^*)^N$. Denote by s the cardinality of S and let $\mathbf{P}$ be the quantity defined in (2.5).

Let

$$\mathcal{X} := \{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \ f_i(\mathbf{x}) = 0, \ i = 1, \ldots, m\}$$

be a subvariety of $(\overline{\mathbb{Q}}^*)^N$, where $f_1, \ldots, f_m$ are non-constant polynomials in $\overline{\mathbb{Q}}[X_1, \ldots, X_N]$ each consisting of 2 or 3 monomials. Put

$$\delta := \max(\deg f_1, \ldots, \deg f_m), \quad H := \max(1, h(f_1), \ldots, h(f_m)).$$

Further, let L be the smallest number field containing K and the coefficients of the polynomials f_i $(i = 1, \ldots, m)$.

The stabilizer of $\mathcal{X}$ is given by

$$\mathrm{Stab}(\mathcal{X}) = \left\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \mid \mathbf{x}\mathcal{X} \subseteq \mathcal{X}\right\},$$

where $\mathbf{x}\mathcal{X} = \{\mathbf{xy} \ : \ \mathbf{y} \in \mathcal{X}\}$. $\mathrm{Stab}(\mathcal{X})$ is clearly an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$, and it can be computed effectively in terms of the polynomials $f_1, \ldots, f_m$ defining $\mathcal{X}$.

Put

$$C^* := \left(e^{11} d^3 r\right)^{r+3} (\delta h_0)^r s \cdot \frac{\mathbf{P}}{\log \mathbf{P}} \cdot \log^* \max(ds\mathbf{P}, \delta h_0), \tag{2.30}$$

and

$$\begin{cases} C_2 := C^* N (2\delta)^{N-1}, \\ C_3 := C^* \cdot 2mh_0 \left(r 4^{r+1} \cdot d(\log 3d)^3 \cdot m\delta h_0\right)^r. \end{cases} \tag{2.31}$$

Theorem 2.11 (Bérczes, Evertse, Győry and Pontreau [11]). *Let $\mathcal{X}$ be a variety satisfying the conditions listed above, and put $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$.*
(i) *Suppose that $\mathcal{H}$ is finite. Then for every $\mathbf{x} \in \mathcal{X} \cap \Gamma$ we have*

$$h(\mathbf{x}) \leq C_2 H.$$

(ii) *Suppose that* $\mathcal{H}$ *is not finite. Then* $\mathcal{X} \cap \Gamma$ *is contained in some finite union of translates*

$$\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H},$$

with

$$\mathbf{x}_i\mathcal{H} \subset \mathcal{X}, \quad \mathbf{x}_i \in \Gamma, \quad h(\mathbf{x}_i) \le C_3 H \text{ for } i = 1, \ldots, T. \tag{2.32}$$

Our results for $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$ and $\mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ are as follows.

Theorem 2.12 (Bérczes, Evertse, Győry and Pontreau [11]). *Put*

$$\varepsilon := \frac{0.03}{4\delta}. \tag{2.33}$$

(i) *Assume that* $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$ *is finite. Then for every* $\mathbf{x} \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$ *we have*

$$h(\mathbf{x}) < rh_0\delta C_2 + C_2 H, \quad [L(\mathbf{x}) : L] \le 2^{m+N}\delta^N. \tag{2.34}$$

(ii) *Assume that* $\mathcal{H}$ *is not finite. Then* $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$ *is contained in a finite union of translates*

$$\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H},$$

where for $i = 1, \ldots, T$, *we have* $\mathbf{x}_i \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$, $\mathbf{x}_i\mathcal{H} \subset \mathcal{X}$, *and where* $h(\mathbf{x}_i)$ *and* $[L(\mathbf{x}_i) : L]$ *are bounded above by effectively computable numbers depending only on* $\Gamma, f_1, \ldots, f_m$.

Remark. It is possible in principle to give explicit expressions for the effectively computable numbers in part (ii) of Theorem 2.12, but these are rather complicated.

Theorem 2.13 (Bérczes, Evertse, Győry and Pontreau [11]). *Let*

$$\varepsilon := \frac{0.03}{4\delta(C_2\delta r h_0 + 2C_2 H)}. \tag{2.35}$$

Assume that $\mathrm{Stab}(\mathcal{X})$ *is finite. Then for every* $\mathbf{x} \in \mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ *we have*

$$h(\mathbf{x}) \le 2rh_0\delta C_2 + 2C_2 H, \quad [L(\mathbf{x}) : L] \le 2^{m+N}\delta^N.$$

Remark. If $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$ is not finite, then in general $\mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ need not be contained in a finite union of translates $\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H}$. Indeed, suppose that $\dim \mathcal{X} > \dim \mathcal{H}$, and that $\mathcal{H} \cap \Gamma$ contains points of infinite order. Pick any $\mathbf{x}_0 \in \mathcal{X}$. Choose a point $\mathbf{u} \in \mathcal{H} \cap \Gamma$ of infinite order. Thus $h(\mathbf{u}) > 0$. Then for any sufficiently large integer n,

$$h(\mathbf{x}_0) \le \varepsilon(1 + nh(\mathbf{u}) - h(\mathbf{x}_0)) \le \varepsilon(1 + h(\mathbf{x}_0\mathbf{u}^n)).$$

Hence $\mathbf{x} := \mathbf{x}_0\mathbf{u}^n \in \mathbf{x}_0\mathcal{H} \cap C(\overline{\Gamma}, \varepsilon)$. That is, every translate $\mathbf{x}_0\mathcal{H}$ with $\mathbf{x}_0 \in \mathcal{X}$ contains elements from $C(\overline{\Gamma}, \varepsilon)$. If $\mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ were contained in a finite union of translates $\cup_{i=1}^t \mathbf{x}_i\mathcal{H}$, then so were $\mathcal{X}$, which is impossible.

Chapter 3

Results over arbitrary finitely generated domains

Let $A := \mathbb{Z}[z_1, \ldots, z_r] \supset \mathbb{Z}$ be a finitely generated integral domain over $\mathbb{Z}$. In this chapter under finitely generated domain we mean an integral domain $\mathbb{Z}[z_1, \ldots, z_r] \supset \mathbb{Z}$, which may contain both algebraic and transcendental elements over $\mathbb{Q}$. Finiteness results for several kinds of Diophantine equations over A date back to the middle of the last century. In his book [45] and paper [44] S. Lang generalized several earlier results on Diophantine equations over the integers to results over A, including results concerning unit equations, Thue-equations and integral points on curves. However, all his results were ineffective. The first effective finiteness results for Diophantine equations over finitely generated domains were published in the 1980's, when Győry [39], [40] developed his new effective specialization method. This enabled him to prove effective results over finitely generated domains of a special type containing also transcendental elements over $\mathbb{Q}$. He proved such results for unit equations, norm form equations, index form equations, discriminant form equations [39] and for polynomials and integral elements of given discriminant [40]. Later Brindza proved such results for superelliptic equations [22] and the generalized Catalan equation [23], Brindza and Pintér obtained such results for equal values of binary forms [24], and Végső [75] for the Schinzel-Tijdeman equation.

In 2013 Evertse and Győry [32] combined the method of Győry with newer results of Aschenbrenner [1] such that they were able to prove effective results for unit equations $ax + by = 1$ in $x, y \in A^*$ over arbitrary finitely generated domains A of characteristic 0.

In this chapter on one hand I present effective finiteness results for Thue equations, hyper- and superelliptic equations, and the Schinzel-Tijdeman equation over arbitrary finitely generated domains. These are joint results with Jan-Hendrik Evertse and Kálmán

Győry published in [10]. On the other hand I present effective finiteness results for generalized unit points and division points on curves over finitely generated domains. These results have been published in my papers [7] and [6].

3.1 Finitely generated domains

Before presenting our results we introduce some concepts and notation.

Let $r > 0$ and let $A := \mathbb{Z}[z_1, \ldots, z_r]$ be a domain of characteristic 0 which is finitely generated over $\mathbb{Z}$. Clearly, A can be expressed as a factor ring

$$A \cong \mathbb{Z}[X_1, \ldots, X_r]/\mathcal{I}, \tag{3.1}$$

where $\mathcal{I}$ is the ideal of $R := \mathbb{Z}[X_1, \ldots, X_r]$ which consists of all polynomials $f \in R$ with the property $f(z_1, \ldots, z_r) = 0$. The ideal $\mathcal{I}$ is finitely generated, so we may write

$$\mathcal{I} = (f_1, \ldots, f_t) \qquad \text{with} \quad f_1, \ldots, f_t \in \mathbb{Z}[X_1, \ldots, X_r]. \tag{3.2}$$

In fact in this way the polynomials $f_1, \ldots, f_t$ fix a representation for the domain A. Recall that A is a domain of characteristic 0 if and only if $\mathcal{I}$ is a prime ideal, and $\mathcal{I} \cap \mathbb{Z} = \emptyset$. Given a set of generators $f_1, \ldots, f_t$ for I this property can be checked effectively (see [1] and [42]).

Let K denote the quotient field of A. We say that the polynomial $f \in R$ *represents* $\alpha \in A$ if we have $f(z_1, \ldots, z_r) = \alpha$. Further we say that the pair $(f, g) \in R^2$ *represents* $\beta \in K$ if $g \notin \mathcal{I}$ (i.e. $g(z_1, \ldots, z_r) \neq 0$) and $\frac{f(z_1,\ldots,z_r)}{g(z_1,\ldots,z_r)} = \beta$. We will also use the terminology that f is a *representative* for α, or (f, g) is a *pair of representatives* for β. Clearly, any element $\alpha \in A$ has infinitely many representatives, and any $\beta \in K$ has infinitely many pairs of representatives. However, since one can effectively decide whether a given polynomial of R belongs to a given ideal of R or not (see [1]), one can also effectively decide if two polynomials represent the same element of A, or if two pairs of polynomials of R represent the same element of K. Indeed, two polynomials $f, f' \in R$ represent the same element $\alpha \in A$ if and only if $f - f' \in \mathcal{I}$, and two pairs of polynomials $(f, g), (f', g') \in R^2$ represent the same element $\beta \in K$ if and only if $fg' - f'g \in \mathcal{I}$.

We shall measure elements of A by their representatives. For a non-zero polynomial $f \in R$ let us denote by $\deg f$ the total degree of f and by $h(f)$ the absolute logarithmic height of f, i.e. the logarithm of the maximum of the absolute values of its coefficients. Further we define the size of f by

$$s(f) := \max(1, \deg f, h(f)).$$

For the constant 0 polynomial we define $s(0) := 1$.

Throughout the dissertation we shall use the notation $O(\cdot)$ to denote a quantity which is c times the expression between the parentheses, where c is an effectively computable positive absolute constant which may be different at each occurrence of the O-symbol. Further, throughout the dissertation we write $\log^* a := \max(1, \log a)$ for $a > 0$, and $\log^* 0 := 1$.

3.2 Effective results for Diophantine equations over finitely generated domains

Let A be an arbitrary integral domain of characteristic 0 that is finitely generated over $\mathbb{Z}$, as defined in Section 3.1. In this section we consider Thue equations $F(x, y) = \delta$ in $x, y \in A$, where F is a binary form with coefficients from A and δ is a non-zero element from A, and hyper- and superelliptic equations $f(x) = \delta y^m$ in $x, y \in A$, where $f \in A[X]$, $\delta \in A \setminus \{0\}$ and $m \in \mathbb{Z}_{\geq 2}$.

Under the necessary finiteness conditions we give effective upper bounds for the sizes of the solutions of the equations in terms of appropriate representations for A, δ, F, f, m. These results imply that the solutions of these equations can be determined in principle. Further, we consider the Schinzel-Tijdeman equation $f(x) = \delta y^m$ where $x, y \in A$ and $m \in \mathbb{Z}_{\geq 2}$ are the unknowns and we give an effective upper bound for m.

We mention that results from the existing literature deal only with equations over restricted classes of finitely generated domains whereas we do not have to impose any restrictions on A. Further, in this generality we give upper bounds for the sizes of the solutions x, y and m that are much more precise than those obtained in the special cases considered earlier.

Our proofs are a combination of existing effective results for Thue equations and hyper- and superelliptic equations over number fields and over function fields, and a recent effective specialization method of Evertse and Győry [32].

3.2.1 Thue equations

Let A be an arbitrary integral domain of characteristic 0 that is finitely generated over $\mathbb{Z}$, as defined in Section 3.1.

We consider the Thue equation over A, i.e. the equation

$$F(x, y) = \delta \qquad \text{in} \quad x, y \in A, \tag{3.3}$$

where

$$F(X,Y) = a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n \in A[X,Y]$$

is a binary form of degree $n \geq 3$ with discriminant $D_F \neq 0$, and $\delta \in A \setminus \{0\}$. We give (3.3) by representatives

$$\tilde{a_0}, \tilde{a_1}, \ldots, \tilde{a_n}, \tilde{\delta} \in \mathbb{Z}[X_1, \ldots, X_r]$$

of $a_0, a_1, \ldots, a_n, \delta$, respectively, where $\delta \notin \mathcal{I}$, and the discriminant $D_{\tilde{F}}$ of the polynomial $\tilde{F} := \sum_{j=0}^{n} \tilde{a_j} X^{n-j}Y^j$ is not in $\mathcal{I}$. These conditions on $\tilde{a_0}, \tilde{a_1}, \ldots, \tilde{a_n}, \tilde{\delta}$ can be checked by means of the ideal membership algorithm mentioned above. Let

$$\begin{cases} \max(\deg f_1, \ldots, \deg f_t, \deg \tilde{a_0}, \deg \tilde{a_1}, \ldots, \deg \tilde{a_n}, \deg \tilde{\delta}) \leq d \\ \max(h(f_1), \ldots, h(f_t), h(\tilde{a_0}), h(\tilde{a_1}), \ldots, h(\tilde{a_n}), h(\tilde{\delta})) \leq h, \end{cases} \tag{3.4}$$

where $d \geq 1$, $h \geq 1$.

Theorem 3.1 (Bérczes, Evertse and Győry [10]). *Every solution x, y of the equation (3.3) has representatives $\tilde{x}, \tilde{y}$ such that*

$$s(\tilde{x}), s(\tilde{y}) \leq \exp\left(n!(nd)^{\exp O(r)}(h+1)\right). \tag{3.5}$$

The exponential dependence of the upper bound on $n!$, d and $h+1$ is coming from a Baker-type effective result for Thue equations over number fields that is used in the proof. The bad dependence on r is coming from the effective commutative algebra for polynomial rings over fields and over $\mathbb{Z}$, that is used in the specialization method of Evertse and Győry mentioned above.

We immediately deduce that equation (3.3) is effectively solvable:

Corollary 3.2 (Bérczes, Evertse and Győry [10]). *There exists an algorithm which, for any given $f_1, \ldots, f_t$ such that A is a domain of characteristic 0, and any representatives $\tilde{a_0}, \ldots, \tilde{a_n}$, $\tilde{\delta}$ such that $D_{\tilde{F}}, \tilde{\delta} \notin \mathcal{I}$, computes a finite list, consisting of one pair of representatives for each solution (x, y) of (3.3).*

Proof. Let C be the upper bound from (3.5). Check for each pair of polynomials $\tilde{x}, \tilde{y} \in \mathbb{Z}[X_1, \ldots, X_r]$ of size at most C whether $\tilde{F}(\tilde{x}, \tilde{y}) - \tilde{\delta} \in \mathcal{I}$. Then for all pairs $\tilde{x}, \tilde{y}$ passing this test, check whether they are equal modulo $\mathcal{I}$, and keep a maximal subset of pairs that are pairwise different modulo $\mathcal{I}$.

3.2.2 Hyper- and superelliptic equations

We now consider the equation

$$F(x) = \delta y^m \qquad \text{in} \quad x, y \in A, \tag{3.6}$$

where

$$F(X) = a_0 X^n + a_1 X^{n-1} + \cdots + a_n \in A[X],$$

$\delta \in A \setminus \{0\}$ and $a_0 \neq 0$, $D_F \neq 0$. Thus, F is a polynomial of degree n without multiple roots. We give (3.6) by representatives

$$\tilde{a_0}, \tilde{a_1}, \ldots, \tilde{a_n}, \tilde{\delta} \in \mathbb{Z}[X_1, \ldots, X_r]$$

for $a_0, a_1, \ldots, a_n, \delta$, respectively, where $\tilde{\delta} \notin \mathcal{I}$, $\tilde{a_0} \notin \mathcal{I}$, and the discriminant of $\tilde{F} := \sum_{j=0}^{n} \tilde{a_j} X^{n-j}$ is not in $\mathcal{I}$. We assume that either $m = 2$ and $n \geq 3$, or $m \geq 3$ and $n \geq 2$. For $m = 2$, equation (3.6) is called a *hyperelliptic equation*, while for $m \geq 3$ it is called a *superelliptic equation*. Let

$$\begin{cases} \max(\deg f_1, \ldots, \deg f_t, \deg \tilde{a_0}, \deg \tilde{a_1}, \ldots, \deg \tilde{a_n}, \deg \tilde{\delta}) \leq d \\ \max(h(f_1), \ldots, h(f_t), h(\tilde{a_0}), h(\tilde{a_1}), \ldots, h(\tilde{a_n}), h(\tilde{\delta})) \leq h, \end{cases} \tag{3.7}$$

where $d \geq 1$, $h \geq 1$.

Theorem 3.3 (Bérczes, Evertse and Győry [10]). *Every solution x, y of the equation (3.6) has representatives $\tilde{x}, \tilde{y}$ such that*

$$s(\tilde{x}), s(\tilde{y}) \leq \exp\left(m^3 (nd)^{\exp O(r)} (h+1)\right). \tag{3.8}$$

Completely similarly as for Thue equations, one can determine effectively a finite list, consisting of one pair of representatives for each solution (x, y) of (3.6).

Our next result deals with the Schinzel-Tijdeman equation, which is (3.6) but with three unknowns $x, y \in A$ and $m \in \mathbb{Z}_{\geq 2}$.

Theorem 3.4 (Bérczes, Evertse and Győry [10]). *Assume that in (3.6) F has non-zero discriminant and $n \geq 2$. Let $x, y \in A$, $m \in \mathbb{Z}_{\geq 2}$ be a solution of (3.6). Then*

$$m \leq \exp\left((nd)^{\exp O(r)} (h+1)\right) \tag{3.9}$$
$$\text{if } y \in \overline{\mathbb{Q}},\ y \neq 0,\ y \text{ is not a root of unity,}$$

$$m \leq (nd)^{\exp O(r)} \quad \text{if } y \notin \overline{\mathbb{Q}}. \tag{3.10}$$

We mention that the condition in (3.9) is necessary, since if y is zero or a root of unity then it is not possible to give upper bound for m.

3.3 Generalized unit points on curves over finitely generated domains

Let A be an arbitrary integral domain of characteristic 0 that is finitely generated over $\mathbb{Z}$, as defined in Section 3.1, let K be its quotient field and A^* its unit group.

Let $F \in A[X, Y]$ be a non-constant polynomial. By a result of Lang [44] from 1960, the equation

$$F(x, y) = 0 \quad \text{in } x, y \in A^* \tag{3.11}$$

has only finitely many solutions, provided F is not divisible by any polynomial of the form

$$X^m Y^n - \alpha \qquad \text{or} \qquad X^m - \alpha Y^n \tag{3.12}$$

for any non-negative integers m, n, not both zero, and any $\alpha \in A^*$. Lang's proof is ineffective. The conditions imposed in Lang's theorem, i.e., that A be finitely generated and F not be divisible by any polynomial of type (3.12), are essentially necessary. On one hand, if A is not finitely generated then it is not possible anymore to prove a finiteness theorem. On the other hand, if F is divisible by a polynomial of the form (3.12), then there exist finitely generated domains A for which equation (3.11) has infinitely many solutions. Furthermore, if F has a divisor of the form (3.12), then if α is a perfect d^{th} power in A, where $d = \gcd(m, n)$, then (3.11) has infinitely many solutions for every domain A which is finitely generated over $\mathbb{Z}$. Bombieri and Gubler [18] (Theorem 5.4.5) gave an effective proof of Lang's result in the case that A is a ring of S-integers in a number field, and this was made more precise, with explicit upper bounds for the heights of x, y, by Bérczes, Evertse, Győry and Pontreau [11].

In this section I present an effective version of Lang's result for arbitrary finitely generated domains A, i.e. I show that given suitable representations for A and the coefficients of F, one can in principle effectively determine the solutions of (3.11) under a slightly stronger condition then (3.12), namely in (3.12) we allow $\alpha \in \overline{K}^*$ instead of $\alpha \in A^*$. In fact, I give a quantitative version of this, with upper bounds for the sizes of x and y. The proof of this result depends on the method developed by Győry [39], [40] and Evertse and Győry [32],

Let A be a finitely generated domain given in the form (3.1), where the ideal $\mathcal{I}$ is generated by the polynomials $f_1, \ldots, f_t \in \mathbb{Z}[X_1, \ldots, X_r]$. Let K denote the quotient field of A and denote by $\overline{K}$ the algebraic closure of K.

Let $I \subset \mathbb{Z}^2_{\geq 0}$ be a non-empty set, and let $F(X, Y) = \sum_{(i,j)\in I} a_{ij} X^i Y^j \in A[X, Y]$ be a

polynomial of total degree $N := \deg F$, and suppose that F fulfils the following condition:

$$\begin{gathered} F \textbf{ is not divisible by} \text{ any non-constant polynomial of the form} \\ X^m Y^n - \alpha \quad \text{or} \quad X^m - \alpha Y^n, \text{where } m, n \in \mathbb{Z}_{\geq 0} \text{ and } \alpha \in \overline{K}^*. \end{gathered} \tag{3.13}$$

Further, suppose that we are given representatives $\tilde{a}_{ij} \in \mathbb{Z}[X_1, \ldots, X_r]$ of $a_{ij} \in A$, respectively. Put $\tilde{F}(X, Y) := \sum_{(i,j)\in I} \tilde{a}_{ij} X^i Y^j$. We assume that

$$\begin{cases} \deg f_1, \ldots, \deg f_t, \deg \tilde{a}_{ij} \leq d \ \text{ for every } \ (i,j) \in I \\ h(f_1), \ldots, h(f_t), h(\tilde{a}_{ij}) \leq h \ \text{ for every } \ (i,j) \in I, \end{cases} \tag{3.14}$$

where d, h are real numbers with $d > 1$ and $h > 1$. In Section 8.1 we show that condition (3.13) is effectively decidable in terms of $f_1, \ldots, f_t$ and the $\tilde{a}_{ij}$.

Theorem 3.5 (Bérczes [7]). *If A is a finitely generated domain as above, and F fulfils the condition (3.13) then for all elements (x, y) of the set*

$$\mathcal{C} := \{(x, y) \in (A^*)^2 | F(x, y) = 0\} \tag{3.15}$$

there exist representatives $\tilde{x}$, $\tilde{y}$, $\tilde{x}'$ and $\tilde{y}'$ of x, y, x^{-1} and y^{-1}, respectively, such that

$$s(\tilde{x}), s(\tilde{y}), s(\tilde{x}'), s(\tilde{y}') \leq \exp\left\{(2d)^{\exp O(r)} (2N)^{(\log^* N)\cdot \exp O(r)} \cdot (h+1)^3\right\}. \tag{3.16}$$

We mention that the above result is effective in the sense that it provides an algorithm to determine, at least in principle, all elements of the set (3.15). Indeed, there are only finitely many polynomials of $\mathbb{Z}[X_1, \ldots, X_r]$ below the bound in (3.16) and these can be effectively enumerated. Further, $(x, y) \in \mathcal{C}$ is clearly fulfilled if and only if there are polynomials $\tilde{x}, \tilde{y}, \tilde{x}', \tilde{y}' \in \mathbb{Z}[X_1, \ldots, X_r]$ with their sizes below the bound (3.16), which fulfil

$$\tilde{x} \cdot \tilde{x}' - 1, \ \tilde{y} \cdot \tilde{y}' - 1, \ \tilde{F}(\tilde{x}, \tilde{y}) \in \mathcal{I}. \tag{3.17}$$

So we can enlist all 4-tuples $(\tilde{x}, \tilde{y}, \tilde{x}', \tilde{y}')$ with $s(\tilde{x}), s(\tilde{y}), s(\tilde{x}'), s(\tilde{y}')$ being smaller than our bound, then (using an ideal membership algorithm) check if (3.17) is fulfilled. Finally, we have to group all the tuples in which $(\tilde{x}, \tilde{y})$ represent the same pair $(x, y) \in (A^*)^2$ and pick out one pair from each group. So we get a list consisting of one representative for each element of the set (3.15).

3.4 Division points on curves over finitely generated domains

Let $A := \mathbb{Z}[z_1, \ldots, z_r] \supset \mathbb{Z}$ be again a finitely generated domain. Let K denote the quotient field of the domain A, and denote by K^* the multiplicative group of non-zero elements of K. Denote by $\overline{K}$ the algebraic closure of K and by $\overline{K}^*$ its unit group. Let Γ be a finitely generated subgroup of K^*. Let $F(X,Y) \in A[X,Y]$ be a polynomial. In 1960 Lang [44] proved that the equation

$$F(x,y) = 0 \quad \text{in } x, y \in \Gamma \tag{3.18}$$

(which is even more general than (3.11)) has only finitely many solutions, provided F is not divisible by any polynomial of the form

$$X^m Y^n - \alpha \qquad \text{or} \qquad X^m - \alpha Y^n \tag{3.19}$$

where m, n are non-negative integers, not both zero, and any $\alpha \in \Gamma$. Lang's proof of this result is ineffective. The first effective versions of this result of Lang have been proved by Bombieri and Gubler [18, p. 147, Theorem 5.4.5] for the number field case (see Bérczes, Evertse Győry and Pontreau [11] or Theorem 2.8 of the present dissertation for an explicit version) and by Bérczes [7] in its full generality, over finitely generated domains (this is Theorem 3.5 of the present dissertation). We mention that the effective results are proved under a slightly stronger condition than (3.19), namely in (3.19) $\alpha \in \overline{K}$ is assumed instead of $\alpha \in \Gamma$.

Denote by $\overline{\Gamma}$ the division group of Γ, i.e. the group defined by

$$\overline{\Gamma} := \left\{ x \in \overline{K}^* \mid \exists\, m \in \mathbb{N},\ x^m \in \Gamma \right\}.$$

Lang also conjectured ([46], [47], see also [48]) that the above equation has finitely many solutions in $x, y \in \overline{\Gamma}$ under the same condition (3.19) but with $\alpha \in \overline{\Gamma}$. Liardet [51], [52] proved this conjecture of Lang. However, this famous result of Liardet is also ineffective.

An effective version of Liardet's Theorem in the number field case is due to Bérczes, Evertse, Győry and Pontreau [11] (see Theorem 2.9 in this dissertation), however, in the general case no effective result had been proved.

In the present section we make effective the above-mentioned finiteness theorem of Liardet in the general case. Our result is not only effective, but also quantitative in the sense that an upper bound for the sizes of the solutions $x, y \in \overline{\Gamma}$ is provided. The presented result is a common generalization of the results of Bombieri and Gubler [18, p. 147, Theorem 5.4.5], Bérczes, Evertse, Győry and Pontreau [11] (see Theorem 2.9 in this

dissertation) and that of Bérczes [7] (see Theorem 3.5 in this dissertation). Further, our result is also a generalization of the result of Bérczes, Evertse and Győry [8] (see Theorem 2.4 in this dissertation) and of Evertse and Győry [32] on unit equations. The main tool of the proof is an effective specialization method introduced by Győry in the 1980's (see [39], [40]), and improved by Evertse and Győry [32] in 2013. The main difficulty of the proof is that on one hand we have to bound also the degrees over K of the solutions from $\overline{\Gamma}$, on the other hand we do not have any convenient representation for the elements of $\overline{\Gamma}$. We also mention that *this is the first effective result for Diophantine equations considered over the division group of an arbitrary finitely generated group.*

Let $A := \mathbb{Z}[z_1, \ldots, z_r]$ be a finitely generated domain over $\mathbb{Z}$, and let K denote its quotient field.

Let $\gamma_1, \ldots, \gamma_s \in K^*$ be arbitrary non-zero elements of K given by corresponding representation pairs $(g_1, h_1), \ldots, (g_s, h_s)$. Define the finitely generated group

$$\Gamma := \left\{ \gamma_1^{l_1} \ldots \gamma_s^{l_s} \mid l_1, \ldots, l_s \in \mathbb{Z} \right\}, \tag{3.20}$$

and its division group

$$\overline{\Gamma} := \left\{ \delta \in \overline{K} \mid \exists\, m \in \mathbb{Z}_{>0} :\ \delta^m \in \Gamma \right\}. \tag{3.21}$$

Let $I \subset \mathbb{Z}_{\geq 0}^2$ be a non-empty set, and let $F(X, Y) = \sum_{(i,j)\in I} a_{ij} X^i Y^j \in A[X, Y]$ be a polynomial which fulfils the following condition:

$$\begin{gathered} F \textbf{ is not divisible by} \text{ any non-constant polynomial of the form} \\ X^m Y^n - \alpha \quad \text{or} \quad X^m - \alpha Y^n, \text{where } m, n \in \mathbb{Z}_{\geq 0} \text{ and } \alpha \in \overline{K}^*. \end{gathered} \tag{3.22}$$

We mention that this condition is effectively decidable, as is explained in Section 8. Let $N := \deg F$ denote the total degree of F, and assume that F is given by specifying a representative $\tilde{a}_{ij}$ of its coefficient a_{ij} for each $(i, j) \in I$.

Further, we assume that

$$\begin{cases} \deg f_1, \ldots, \deg f_t, \deg g_1, \ldots, \deg g_s, \deg h_1, \ldots, \deg h_s, \deg \tilde{a}_{ij} \leq d \\ h(f_1), \ldots, h(f_t), h(g_1), \ldots, h(g_s), h(h_1), \ldots, h(h_s), h(\tilde{a}_{ij}) \leq h, \end{cases} \tag{3.23}$$

where $(i, j) \in I$ and d, h are real numbers with $d > 1$ and $h > 1$.

Theorem 3.6 (Bérczes [6]). *Let A be a finitely generated domain as above, $\overline{\Gamma}$ the above-defined division group and $F(X, Y) \in A[X, Y]$ a polynomial which fulfils the condition (3.22). Define the set*

$$\mathcal{C} := \{(x, y) \in (\overline{\Gamma})^2 | F(x, y) = 0\}. \tag{3.24}$$

(i) *Then there exists a positive integer m with*

$$m \leq \exp\left\{N^6(2d)^{\exp\{C_3(r+s)\}}(h+1)^{4s}\right\} \tag{3.25}$$

with C_3 an effectively computable absolute constant such that

$$x^m \in \Gamma \qquad \text{and} \qquad y^m \in \Gamma, \qquad \text{for every } (x,y) \in \mathcal{C}.$$

(ii) *More precisely, there exists an effectively computable absolute constant C_4, such that for all $(x,y) \in \mathcal{C}$ there are integers $t_{1,x}, \ldots, t_{s,x}, t_{1,y}, \ldots, t_{s,y}$ with*

$$t_{i,x}, t_{i,y} \leq \exp\left\{\exp\left\{N^{12}(2d)^{\exp\{C_4(r+s)\}}(h+1)^{8s}\right\}\right\} \tag{3.26}$$

for $i = 1, \ldots, s$, such that

$$x^m = \gamma_1^{t_{1,x}} \ldots \gamma_s^{t_{s,x}}, \qquad y^m = \gamma_1^{t_{1,y}} \ldots \gamma_s^{t_{s,y}}. \tag{3.27}$$

Part II

Proofs

Chapter 4

Proof of the results from Section 2.2

4.1 Proof of Theorems 2.6 and 2.7

We need several auxiliary results.

Keeping the notation of Section 2.2, let K be an algebraic number field of degree d and assume that it is embedded in $\mathbb{C}$. Let

$$\Lambda = \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1, \tag{4.1}$$

where $\alpha_1, \ldots, \alpha_n$ are n (≥ 2) non-zero elements of K, and $b_1, \ldots, b_n$ are rational integers, not all zero. Put

$$B^* = \max\{|b_1|, \ldots, |b_n|\}.$$

Let $A_1, \ldots, A_n$ be reals with

$$A_i \geq \max\{dh(\alpha_i), \pi\} \quad (i = 1, \ldots, n). \tag{4.2}$$

Theorem A. (Matveev [56]) *Let $n \geq 2$. Suppose that $\Lambda \neq 0$, $b_n = \pm 1$, and let B be a real number with*

$$B \geq \max\left\{B^*, 2e \max\left(\frac{n\pi}{\sqrt{2}}, A_1, \ldots, A_{n-1}\right) A_n\right\}. \tag{4.3}$$

Then we have

$$\log|\Lambda| > -c_1(n,d) A_1 \cdots A_n \log(B/(\sqrt{2}A_n)), \tag{4.4}$$

where

$$c_1(n,d) = \min\left\{1.451(30\sqrt{2})^{n+4}(n+1)^{5.5}, \pi 2^{6.5n+27}\right\} d^2 \log(ed).$$

Proof. This is a consequence of Corollary 2.3 of Matveev [56]; see Proposition 4 in Győry and Yu [41]. □

Let B and B_n be real numbers satisfying

$$B \geq \max\{|b_1|, \ldots, |b_n|\}, \quad B \geq B_n \geq |b_n|. \tag{4.5}$$

Denote by $\mathfrak{p}$ a prime ideal of the ring of integers $\mathcal{O}_K$ and let $e_\mathfrak{p}$ and $f_\mathfrak{p}$ be the ramification index and the residue class degree of $\mathfrak{p}$, respectively. Thus $N(\mathfrak{p}) = p^{f_\mathfrak{p}}$, where p is the prime number below $\mathfrak{p}$.

Theorem B. (Yu [78]) *Let $n \geq 2$. Assume that $\operatorname{ord}_p b_n \leq \operatorname{ord}_p b_i$ for $i = 1, \ldots, n$, and set*

$$h_i' = \max\{h(\alpha_i), 1/(16e^2d^2)\}, \quad i = 1, \ldots, n.$$

If $\Lambda \neq 0$, then for any real δ with $0 < \delta \leq 1/2$ we have

$$\begin{aligned} \operatorname{ord}_\mathfrak{p} \Lambda \leq c_2(n,d) e_\mathfrak{p}^n \frac{N(\mathfrak{p})}{(\log N(\mathfrak{p}))^2} \cdot \\ \cdot \max\left\{ h_1' \cdots h_n' \log(M\delta^{-1}), \frac{\delta B}{B_n c_3(n,d)} \right\}, \end{aligned} \tag{4.6}$$

where

$$\begin{aligned} c_2(n,d) &= (16ed)^{2(n+1)} n^{3/2} \log(2nd) \log(2d), \\ c_3(n,d) &= (2d)^{2n+1} \log(2d) \log^3(3d), \end{aligned}$$

and

$$M = B_n c_4(n,d) N(\mathfrak{p})^{n+1} h_1' \ldots h_{n-1}',$$

with

$$c_4(n,d) = 2e^{(n+1)(6n+5)} d^{3n} \log(2d).$$

Proof. This is the second consequence of the Main Theorem in Yu [78]. □

The following theorem is a consequence of Theorems A and B.

Theorem C. *Let $n \geq 2$ and $v \in M_K$. Suppose that in (4.1) we have $\Lambda \neq 0$, $b_n = \pm 1$ and that $\alpha_1, \ldots, \alpha_{n-1}$ are not roots of unity. Let*

$$Q_\alpha := h(\alpha_1) \cdots h(\alpha_{n-1}), \quad H := \max(h(\alpha_n), 1).$$

If

$$B \geq \max(|b_1|, \ldots, |b_{n-1}|, 2e(3d)^{2n} Q_\alpha H), \tag{4.7}$$

then

$$\log |\Lambda|_v > -c_5(n,d) \frac{P(v)}{\log P(v)} Q_\alpha H \log^* \left(\frac{BP(v)}{H} \right), \tag{4.8}$$

where $P(v)$ is defined in (2.4) and

$$c_5(n,d) = \lambda(16ed)^{3n+2}(\log^* d)^2,$$

with $\lambda = 1$ or 12 according as $n \geq 3$ or $n = 2$.

To deduce Theorem C from Theorems A and B, we need the following.

Lemma 4.1. (Voutier [76]) *Suppose that α is a non-zero algebraic number of degree d which is not a root of unity. Then*

$$dh(\alpha) \geq \begin{cases} \log 2 & \text{if } d = 1, \\ 2/(\log 3d)^3 & \text{if } d \geq 2. \end{cases} \tag{4.9}$$

Proof. For $d \geq 2$ this is due to Voutier [76]. He showed also that for $d \geq 2$ this lower bound may be replaced by $(1/4)(\log\log d/\log d)^3$. □

Proof of Theorem C. First assume that v is infinite. We apply Theorem A with $A_i = \max\{dh(\alpha_i), \pi\}$ for $i = 1, \ldots, n$. Then using (4.9), it is easy to see that

$$A_1 \cdots A_n \leq (2.52d)^{2n} Q_\alpha H.$$

Further, we have $\sqrt{2}A_n > H/P(v)$ and

$$2e \max\left\{\frac{n\pi}{\sqrt{2}}, A_1, \ldots, A_{n-1}\right\} A_n \leq 2e(3d)^{2n} Q_\alpha H.$$

Now (4.7) implies (4.3), and (4.8) follows from the inequality (4.4) of Theorem A.

Next assume that v is finite. Keeping the notation of Theorem B and using again (4.9), we infer that

$$h_i' = h(\alpha_i) \text{ for } i = 1, \ldots, n-1 \qquad h_n' = h(\alpha_n).$$

Hence $h_n' = H$ if $h(\alpha_n) \geq 1$ and $H = 1$ otherwise. Choosing $\delta = h_1' \cdots h_n'/B$ and $B_n = 1$ in Theorem B, (4.7) implies that $\delta \leq \frac{1}{2}$. Using the fact that $|\Lambda|_v = N(\mathfrak{p})^{-\operatorname{ord}_\mathfrak{p} \Lambda}$, after some computation (4.8) follows from (4.6) of Theorem B. □

Theorem 2.6 will be proved by combining Theorem C with the following result from the geometry of numbers. Let t be a positive integer. A convex distance function on $\mathbb{R}^t$ is a function $f := \mathbb{R}^t \to \mathbb{R}_{\geq 0}$ such that

$$\begin{aligned} &f(\mathbf{x}+\mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^t, \\ &f(\lambda\mathbf{x}) = |\lambda| f(\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathbb{R}^t, \lambda \in \mathbb{R}, \\ &f(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}. \end{aligned}$$

Lemma 4.2. *Let f be a convex distance function on $\mathbb{R}^t$. Let $\{\mathbf{a}_1, \dots, \mathbf{a}_t\}$ be any basis of $\mathbb{Z}^t$ for which the product $f(\mathbf{a}_1) \cdots f(\mathbf{a}_t)$ is minimal. Let $\mathbf{x} \in \mathbb{Z}^t$ and suppose that $\mathbf{x} = b_1\mathbf{a}_1 + \cdots + b_t\mathbf{a}_t$ with $b_1, \dots, b_t \in \mathbb{Z}$. Then*

$$\max(|b_1|f(\mathbf{a}_1), \dots, |b_t|f(\mathbf{a}_t)) \le c_6(t)f(\mathbf{x}), \tag{4.10}$$

where $c_6(t) = t^{2t}$.

Remark. Schlickewei [66] proved that there exists a basis $\{\mathbf{a}_1, \dots, \mathbf{a}_t\}$ of $\mathbb{Z}^t$ satisfying (4.10) with 4^t instead of $c_6(t)$, but it is not clear whether for this basis, the product $f(\mathbf{a}_1) \cdots f(\mathbf{a}_t)$ is minimal. In our proof of Theorem 2.6, the minimality of $f(\mathbf{a}_1) \cdots f(\mathbf{a}_t)$ is crucial, while an improvement of $c_6(t)$ would have only little influence on the final result.

Proof. Let $C = \{\mathbf{x} \in \mathbb{R}^t : f(\mathbf{x}) \le 1\}$. This is a compact, convex body which is symmetric around $\mathbf{0}$. Let $\lambda_1, \dots, \lambda_t$ denote the successive minima of C with respect to the lattice $\mathbb{Z}^t$. Since $\lambda_1 \le \cdots \le \lambda_t$, it follows from a result of Mahler (see e.g. Cassels [28], pp. 135-136, Lemma 8) that there exists a basis $\mathbf{y}_1, \dots, \mathbf{y}_t$ of $\mathbb{Z}^t$ such that $f(\mathbf{y}_i) \le \max(1, i/2)\lambda_i$. Together with Minkowski's theorem on successive minima, this gives

$$f(\mathbf{a}_1) \cdots f(\mathbf{a}_t) \le f(\mathbf{y}_1) \cdots f(\mathbf{y}_t) \le 2t! \cdot \mathrm{Vol}(C)^{-1}, \tag{4.11}$$

where $\mathrm{Vol}(C)$ denotes the volume of C.

By Jordan's theorem or John's Lemma (see e.g. Schmidt [68], pp. 87–89) there is a t-dimensional ellipsoid E in $\mathbb{R}^t$ such that $E \subseteq C \subseteq (\sqrt{t})E$. Further, there is a $t \times t$ real non-singular matrix A such that $E = \{\mathbf{x} \in \mathbb{R}^t \ : \ \|A\mathbf{x}\| \le 1\}$, where $\|\cdot\|$ denotes the Euclidean norm. Thus

$$\frac{1}{\sqrt{t}}\|A\mathbf{x}\| \le f(\mathbf{x}) \le \|A\mathbf{x}\| \quad \text{for} \ \ \mathbf{x} \in \mathbb{R}^t. \tag{4.12}$$

Consequently,

$$V(t)|\det(A)|^{-1} \le \mathrm{Vol}(C) \le t^{t/2}V(t)|\det(A)|^{-1}, \tag{4.13}$$

where $V(t)$ denotes the volume of the t-dimensional unit ball.

Now let $\mathbf{x} = b_1\mathbf{a}_1 + \cdots + b_t\mathbf{a}_t$ with $b_1, \dots, b_t \in \mathbb{Z}$. Then $A\mathbf{x} = b_1(A\mathbf{a}_1) + \cdots + b_t(A\mathbf{a}_t)$. Let B be the matrix with columns $A\mathbf{a}_1, \dots, A\mathbf{a}_t$. Since $|\det(\mathbf{a}_1, \dots, \mathbf{a}_t)| = 1$, we have $|\det(B)| = |\det(A)|$. By this fact, Cramer's rule and Hadamard's inequality, we have for $i = 1, \dots, t$,

$$\begin{aligned} |b_i| &= |\det(A\mathbf{a}_1, \dots, A\mathbf{a}_{i-1}, A\mathbf{x}, A\mathbf{a}_{i+1}, \dots A\mathbf{a}_t)| \ \big/ \ |\det(B)| \\ &\le \|A\mathbf{a}_1\| \cdots \|A\mathbf{a}_{i-1}\| \cdot \|A\mathbf{x}\| \cdot \|A\mathbf{a}_{i+1}\| \cdots \|A\mathbf{a}_t\| \ \big/ \ |\det(A)|. \end{aligned}$$

Together with (4.12), (4.11) and (4.13), this implies

$$\begin{aligned} |b_i| f(\mathbf{a}_i) &\le t^{(t-1)/2}\big(f(\mathbf{a}_1)\cdots f(\mathbf{a}_t)/|\det(A)|\big) f(\mathbf{x}) \\ &\le t^{(t-1)/2}\cdot 2t! V(t)^{-1} f(\mathbf{x}) \qquad \text{for } \ i=1,\dots,t. \end{aligned}$$

By inserting $V(t)=\pi^{t/2}/(t/2)!$ if t is even and $V(t)=\pi^{(t-1)/2}\big/\big(\frac{1}{2}\cdot\frac{3}{2}\cdots\frac{t}{2}\big)$ if t is odd, we get the bound in (4.10). □

Lemma 4.3. *Let G be a finitely generated multiplicative subgroup of K^* of rank $t>0$. Let $\delta_1,\dots,\delta_t\in G$ be multiplicatively independent such that $h(\delta_1)\le\cdots\le h(\delta_t)$. Then G/G_{tors} has a basis $\{\gamma_1,\dots,\gamma_t\}$ such that*

$$h(\gamma_i)\le \max(1,i/2)h(\delta_i) \quad \textit{for } \ i=1,\dots,t. \tag{4.14}$$

Proof. Let $\{\rho_1,\dots,\rho_t\}$ be a basis for $G/G_{\rm tors}$. Then we can write

$$\delta_i=\zeta_i\rho_1^{b_{i1}}\cdots\rho_t^{b_{it}}, \quad i=1,\dots,t,$$

where $\zeta_i\in G_{\rm tors}$, and

$$\mathbf{b}_1=(b_{11},\dots,b_{1t}),\dots,\mathbf{b}_t=(b_{t1},\dots,b_{tt})$$

are linearly independent vectors in $\mathbb{Z}^t$.

Let $S\subset M_K$ be minimal such that S contains all infinite places and $G\subseteq\mathcal{O}_S^*$. We define

$$f(\mathbf{x}):=\frac{1}{2}\sum_{v\in S}\left|x_1\log|\rho_1|_v+\cdots+x_t\log|\rho_t|_v\right|,$$

where $\mathbf{x}=(x_1,\dots,x_t)\in\mathbb{R}^t$. This is a convex distance function. Further, by (2.6) we have

$$f(\mathbf{b}_i)=h(\delta_i) \quad \text{for } \ i=1,\dots,t. \tag{4.15}$$

Using again Mahler's result mentioned above, we infer that there is a basis $\mathbf{a}_i=(a_{i1},\dots,a_{it})$ $(i=1,\dots,t)$ of $\mathbb{Z}^t$ for which

$$f(\mathbf{a}_i)\le\max(1,i/2)f(\mathbf{b}_i) \quad \text{for } \ i=1,\dots,t. \tag{4.16}$$

Putting $\gamma_i=\rho_1^{a_{i1}}\cdots\rho_t^{a_{it}}$ for $i=1,\dots,t$, we infer that $\{\gamma_1,\dots,\gamma_t\}$ is a basis for $G/G_{\rm tors}$, which in view of (4.15), (4.16) satisfies (4.14). □

We first prove Theorem 2.6 and then Theorem 2.7.

Proof of Theorem 2.6. Since G has rank $t > 0$, there are t multiplicatively independent elements among the generators $\xi_1, \dots, \xi_r$, say $\xi_1, \dots, \xi_t$. Then by Lemma 4.1

$$h(\xi_1) \cdots h(\xi_t) \le c_7(d)^{r-t} Q_G, \tag{4.17}$$

where $c_7(d) = \frac{d}{2}(\log 3d)^3$ if $d \ge 2$ and $c_7(d) = (\log 2)^{-1}$ if $d = 1$. Let $\delta_1, \dots, \delta_t$ be multiplicatively independent elements of G such that $h(\delta_1) \cdots h(\delta_t)$ is minimal. Then

$$h(\delta_1) \cdots h(\delta_t) \le h(\xi_1) \cdots h(\xi_t). \tag{4.18}$$

Further, by Lemma 4.3, G/G_{tors} has a basis $\{\gamma_1, \dots, \gamma_t\}$ such that

$$h(\gamma_1) \cdots h(\gamma_t) \le c_8(t) h(\delta_1) \cdots h(\delta_t), \tag{4.19}$$

with $c_8(t) := t!/2^{t-1}$. We may assume that $\{\gamma_1, \dots, \gamma_t\}$ is such a basis of G/G_{tors} for which $h(\gamma_1) \cdots h(\gamma_t)$ is minimal.

For $\xi \in G$, we can write

$$\xi = \zeta \gamma_1^{b_1} \cdots \gamma_t^{b_t}, \tag{4.20}$$

where $\zeta \in G_{\text{tors}}$ and $\mathbf{b} = (b_1, \dots, b_t) \in \mathbb{Z}^t$. As in the proof of Lemma 4.3, consider the following convex distance function on $\mathbb{R}^t$:

$$f(\mathbf{x}) := \frac{1}{2} \sum_{v \in S} \left| x_1 \log |\gamma_1|_v + \cdots + x_t \log |\gamma_t|_v \right|,$$

where $\mathbf{x} = (x_1, \dots, x_t) \in \mathbb{R}^t$ and S is the same as in the proof of Lemma 4.3. Then $f(\mathbf{b}) = h(\xi)$. Consider the standard basis $\mathbf{a}_1 = (1, 0, \dots, 0)$, $\mathbf{a}_2 = (0, 1, 0, \dots, 0)$, …, $\mathbf{a}_t = (0, \dots, 0, 1)$ in $\mathbb{Z}^t$. Then

$$f(\mathbf{a}_i) = h(\xi_i) \quad \text{for} \quad i = 1, \dots, t,$$

and $f(\mathbf{a}_1) \cdots f(\mathbf{a}_t)$ is minimal among the bases of $\mathbb{Z}^t$.

We can now apply Lemma 4.2 to this basis $\mathbf{a}_1, \dots, \mathbf{a}_t$, and infer that

$$|b_i| h(\gamma_i) = |b_i| f(\mathbf{a}_i) \le c_6(t) f(\mathbf{b}) = c_6(t) h(\xi), \quad i = 1, \dots, t.$$

Together with Lemma 4.1 this gives

$$\max(|b_1|, \dots, |b_t|) \le c_6(t) c_7(d) h(\xi). \tag{4.21}$$

We apply now Theorem C with $v \in S$ and with

$$\Lambda = 1 - \alpha\xi = 1 - \alpha' \gamma_1^{b_1} \cdots \gamma_t^{b_t},$$

where $\alpha' = \zeta\alpha$. Let $Q_\gamma := h(\gamma_1)\cdots h(\gamma_t)$. First assume that

$$c_6(t)c_7(d)h(\xi) \geq 2e(3d)^{2(t+1)}Q_\gamma H. \tag{4.22}$$

Further suppose that

$$h(\xi) \geq (c_6(t)c_7(d))^{1/2}\, H. \tag{4.23}$$

Then putting $B = c_6(t)c_7(d)h(\xi)$, it follows that

$$\log^*\left(\frac{BP(v)}{H}\right) \leq 3\log\left(\frac{h(\xi)P(v)}{H}\right). \tag{4.24}$$

Together with (4.17), (4.18), (4.19) and (4.24), Theorem C gives (2.24) after some computation.

Consider now the case when at least one of (4.22) and (4.23) does not hold. We cover this remaining case by assuming that

$$h(\xi) < \frac{1}{2}c_2(r,d,t)Q_G H$$

with the $c_2(r,d,t)$ occurring in Theorem 2.6. By the product formula and Liouville's inequality we get

$$\begin{aligned}|1-\alpha\xi|_v = \prod_{\substack{w\in M_K\\ w\neq v}} |1-\alpha\xi|_w^{-1} &\geq \frac{1}{2}\prod_{\substack{w\in M_K\\ w\neq v}} \max(1,|\alpha\xi|_w)^{-1}\\ &\geq \frac{1}{2}\exp(-h(\alpha\xi)) \geq \frac{1}{2}\exp\left(-\left(H+\frac{1}{2}c_2(r,d,t)Q_G H\right)\right),\end{aligned}$$

whence (2.24) follows again.

Finally, assume that $r = t$ and that $\{\xi_1,\ldots,\xi_t\}$ is a basis of G/G_{tors}. We may assume without loss of generality that $Q_G = h(\xi_1)\cdots h(\xi_t)$ is minimal among all bases of G/G_{tors}. Then in our above proof we can choose $\gamma_i = \xi_i$ for $i = 1,\ldots,t$ and we do not need $\delta_1,\ldots,\delta_t$. This simplification in the proof gives (2.24) with $c_2(d,t)$ in place of $c_2(r,d,t)$. □

Proof of Theorem 2.7. Together with the estimate (2.24) of Theorem 2.6, (2.25) gives (2.26), and then (2.27) easily follows. □

4.2 Proof of Theorems 2.1, 2.2, 2.3 and 2.5

Taking as a starting point Theorem 2.7, we first deduce Theorem 2.2, then Theorem 2.1, and from the latter Theorems 2.3 and 2.5.

Proof of Theorem 2.2. First suppose that $a_1, a_2 \in K$. Let (x_1, x_2) be a solution of (2.10). Then (2.10) gives

$$h(x_1) \le 3H + h(x_2) + \log 2. \tag{4.25}$$

First assume that $h(x_2) < 4 \cdot 10^2 sH$. Then (4.25) gives $h(x_1) \le 404sH$, and by this we get $h(x_1)\mathbf{P}/H \le 404s\mathbf{P}$. Using now the fact that the function $X/\log X$ is monotone increasing for $X > e$, (2.11) and (2.12) easily follow.

Now assume that

$$h(x_2) \ge 4 \cdot 10^2 sH. \tag{4.26}$$

Choose $v \in S$ for which $|x_2|_v$ is minimal. Then we infer from (2.10) that

$$\log|1 - a_1x_1|_v = \log|a_2x_2|_v \le -\frac{1}{s}h(x_2) + H. \tag{4.27}$$

Further, it follows from (4.25) and (4.26) that $h(x_1) \le 1.01h(x_2)$. Hence we get from (4.26) and (4.27) that

$$\log|1 - a_1x_1|_v < -\kappa h(x_1)$$

with the choice $\kappa = 1/(2.02s)$. By applying the estimate (2.26) of Theorem 2.7 we deduce (2.11) and subsequently we get for $h(x_1)$ the upper bound in (2.12) with 6.5 replaced by 6.4. Finally, it follows from (2.10) that $h(x_2) \le 3H + h(x_1) + \log 2$, so we obtain (2.12) for $h(x_2)$ as well.

Now suppose that $(a_1, a_2) \notin (K^*)^2$. Then we choose a nontrivial embedding σ of the extension K_0/K into $\mathbb{C}$, where $K_0 = K(a_1, a_2)$. Then equation (2.10) leads to

$$\sigma(a_1)x_1 + \sigma(a_2)x_2 = 1. \tag{4.28}$$

Now expressing x_1 and x_2 by Cramer's rule from the system consisting of (2.10) and (4.28) we get an estimate for $h(x_1)$ and $h(x_2)$ which is much sharper than (2.11) and (2.12). □

Proof of Theorem 2.1. Suppose that $\xi_1, \dots, \xi_r$ generate a multiplicative subgroup, say G, of $\overline{\mathbb{Q}}^*$ of rank $t > 0$. Clearly G is contained in K^*. We may assume that $\xi_1, \dots, \xi_{r'}$ are not roots of unity. Then $t \le r' \le r$ and $\xi_1, \dots, \xi_{r'}$ is a system of generators of G/G_{tors}. By the assumption made on $\mathbf{w}_1, \dots, \mathbf{w}_r$, $\eta_{r'+1}, \dots, \eta_r$ are not roots of unity. Put

$$Q_G := h(\xi_1) \cdots h(\xi_{r'}).$$

Using Lemma 4.1 we infer that

$$Q_G \le c_7(d)^{r-r'} Q_\Gamma, \tag{4.29}$$

where $c_7(d) = (1/2)d(\log 3d)^3$ if $d \ge 2$ and $c_7(d) = (\log 2)^{-1}$ if $d = 1$.

Let (x_1, x_2) be a solution of (2.7). Then $x_1 \in G$ and $x_2 \in \mathcal{O}_S^*$. We can now apply Theorem 2.2 to this solution and we obtain (2.12) with r replaced by r'. Using $r' \leq r$, (4.29) and

$$h(x_1, x_2) \leq h(x_1) + h(x_2),$$

(2.9) easily follows from (2.12). □

In the proofs of Theorems 2.3 and 2.5 we need the following lemma.

Lemma 4.4. (Beukers and Zagier). *Let* $(b_1, b_2) \in (\overline{\mathbb{Q}}^*)^2$, *and let* (x_{i1}, x_{i2}) $(i = 1, 2, 3)$ *be points in* $(\overline{\mathbb{Q}}^*)^2$ *with* $b_1 x_{i1} + b_2 x_{i2} = 1$ *for* $i = 1, 2, 3$. *Then we have*

$$\sum_{i=1}^{3} h(x_{i1}, x_{i2}) \geq 0.09. \tag{4.30}$$

Proof. By Corollary 2.4 in [14] we have $\sum_{i=1}^{3} h(x_{i1}, x_{i2}) \geq \log \rho$, where ρ denotes the real root of $\rho^{-6} + \frac{1}{2}\rho^{-2} = 1$ which is larger than 1. We have $\log \rho \geq 0.09$. □

The proofs of Theorems 2.3 and 2.5 are very similar. We work out the proof of Theorem 2.5 in detail, and then indicate which changes have to be made to obtain Theorem 2.3.

Proof of Theorem 2.5. Fix a solution (x_1, x_2) of equation (2.20). Since $(x_1, x_2) \in C(\overline{\Gamma}, \varepsilon)$ we can write

$$\begin{aligned} &(x_1, x_2) = (y_1, y_2)(z_1, z_2) \ \text{ with} \\ &(y_1, y_2) \in \overline{\Gamma}, \quad h(z_1, z_2) < \varepsilon(1 + h(y_1, y_2)). \end{aligned} \tag{4.31}$$

Further, we can write

$$\begin{aligned} &(y_1, y_2) = (y_1', y_2')(w_1, w_2) \ \text{ with} \\ &(y_1', y_2') \in \Gamma, \\ &(w_1, w_2) = \prod_{i=1}^{r} (\xi_i, \eta_i)^{\gamma_i} \text{ with } \gamma_i \in \mathbb{Q},\ |\gamma_i| \leq \frac{1}{2} \ \ (i = 1, \ldots, r). \end{aligned} \tag{4.32}$$

(Note that w_1, w_2 are defined up to roots of unity.) Thus we have

$$h(w_1, w_2) \leq \sum_{i=1}^{r} |\gamma_i| h(\xi_i, \eta_i) \leq r h_0. \tag{4.33}$$

Write

$$(a_1', a_2') := (a_1, a_2)(w_1, w_2)(z_1, z_2). \tag{4.34}$$

Then by (4.33), (4.31),

$$h(a_1', a_2') \le h(a_1, a_2) + rh_0 + \varepsilon(1 + h(y_1, y_2))$$

which leads to

$$h(a_1', a_2') \le h(a_1, a_2) + rh_0 + \varepsilon(1 + h(y_1', y_2') + rh_0). \tag{4.35}$$

Further, equation (2.20) can be written in the form

$$a_1'y_1' + a_2'y_2' = 1 \qquad \text{in } (y_1', y_2') \in \Gamma. \tag{4.36}$$

Using Theorem 2.1 we get

$$h(y_1', y_2') \le A \max\{h(a_1', a_2'), 1\} \tag{4.37}$$

where A is the constant defined in (2.8). Notice that this constant does not depend on the field generated by a_1', a_2'. Further, using (4.35) we get

$$h(y_1', y_2') \le Ah(a_1, a_2) + rh_0A + \varepsilon A + \varepsilon Ah(y_1', y_2') + rh_0\varepsilon A.$$

Since in view of (2.21) we have $\varepsilon < \frac{1}{2A}$ we obtain

$$h(y_1', y_2') \le 2Ah(a_1, a_2) + (1 + 2Arh_0 + rh_0). \tag{4.38}$$

Now by

$$h(y_1, y_2) \le h(y_1', y_2') + h(w_1, w_2) \le 2Ah(a_1, a_2) + (1 + 2Arh_0 + 2rh_0) \tag{4.39}$$

and

$$h(x_1, x_2) \le h(y_1, y_2) + \varepsilon(1 + h(y_1, y_2)) \le (\varepsilon + 1)h(y_1, y_2) + \varepsilon$$

we get

$$h(x_1, x_2) \le 3Ah(a_1, a_2) + 5Arh_0 \tag{4.40}$$

which proves assertion (2.22) of our Theorem 2.5.

Now we have to prove the explicit upper bound on $[K_0(x_1, x_2) : K_0]$, where (x_1, x_2) is any solution of (2.20) and K_0 is the field generated by Γ, a_1, a_2. Let us fix such a solution. Choose (y_1, y_2), (z_1, z_2) as in (4.31) and then (y_1', y_2'), (w_1, w_2) as in (4.32). Finally, define (a_1', a_2') by (4.34). Define the field $L := K_0(a_1', a_2')$. We first prove that $[L : K_0] \le 2$.

Assume that this is false, that is, $[L : K_0] \ge 3$. Then there are at least 3 distinct embeddings of L to $\mathbb{C}$ which leave fixed the field K_0, call them $\sigma_1, \sigma_2, \sigma_3$. We consider again equation (4.36). Since $(y_1', y_2') \in \Gamma \subset (K_0^*)^2$ we have

$$\sigma_i(a_1')y_1' + \sigma_i(a_2')y_2' = 1 \quad \text{for } i = 1, 2, 3.$$

This means that the equation

$$(a_1'y_1')X + (a_2'y_2')Y = 1 \qquad \text{in } (X,Y) \in (\overline{\mathbb{Q}}^*)^2$$

has at least 3 distinct solutions, namely $\left(\frac{\sigma_i(a_1')}{a_1'}, \frac{\sigma_i(a_2')}{a_2'}\right)$ $(i = 1, 2, 3)$. Now using Lemma 4.4 we know that

$$\sum_{i=1}^{3} h\left(\frac{\sigma_i(a_1')}{a_1'}, \frac{\sigma_i(a_2')}{a_2'}\right) \geq 0.09. \tag{4.41}$$

On the other hand by (4.34) we have for any embedding $\sigma : L \to \mathbb{C}$,

$$\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) = \left(\frac{\sigma(a_1)}{a_1}, \frac{\sigma(a_2)}{a_2}\right)\left(\frac{\sigma(w_1)}{w_1}, \frac{\sigma(w_2)}{w_2}\right)\left(\frac{\sigma(z_1)}{z_1}, \frac{\sigma(z_2)}{z_2}\right).$$

However, $a_1, a_2 \in K_0$. Further, $(w_1, w_2) \in \overline{\Gamma}$, hence there exists a positive integer m such that $(w_1, w_2)^m \in \Gamma$. This means that $\left(\frac{\sigma(w_1)}{w_1}\right)^m = 1$ and $\left(\frac{\sigma(w_2)}{w_2}\right)^m = 1$. Thus we see that there exist roots of unity ζ_1, ζ_2 such that $\sigma(w_1) = \zeta_1 w_1$ and $\sigma(w_2) = \zeta_2 w_2$. So

$$\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) = (\zeta_1, \zeta_2)\left(\frac{\sigma(z_1)}{z_1}, \frac{\sigma(z_2)}{z_2}\right)$$

and together with $(x_1, x_2) \in C(\overline{\Gamma}, \varepsilon)$ and (4.39),

$$\begin{aligned} h\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) &\leq 2h(z_1, z_2) \leq 2\varepsilon(1 + h(y_1, y_2))) \\ &\leq 2\varepsilon(2Ah(a_1, a_2) + (1 + 2Arh_0 + 2rh_0)) \\ &\leq 2\varepsilon(2Ah(a_1, a_2) + 5Arh_0). \end{aligned} \tag{4.42}$$

This shows that

$$\sum_{i=1}^{3} h\left(\frac{\sigma_i(a_1')}{a_1'}, \frac{\sigma_i(a_2')}{a_2'}\right) < 4\varepsilon(2Ah(a_1, a_2) + 5Arh_0) < 0.09, \tag{4.43}$$

and this contradicts (4.41). Thus, we have proved that $[L : K_0] \leq 2$.

In view of $y_1', y_2' \in K_0$ this shows that $[K_0(a_1'y_1', a_2'y_2') : K_0] \leq 2$, which in turn means that $[K_0(a_1x_1, a_2x_2) : K_0] \leq 2$ and finally, using that $a_1, a_2 \in K_0$ we get $[K_0(x_1, x_2) : K_0] \leq 2$.

□

Proof of Theorem 2.3. The proof of Theorem 2.3 is completely similar to the proof of Theorem 2.5. The only difference is that the estimate (4.31) for $h(z_1, z_2)$ has to be

replaced by $h(z_1, z_2) < \varepsilon$. This slightly modifies the estimates in the proof of Theorem 2.5 and instead of (4.38) we get

$$h(y_1', y_2') \leq Ah(a_1, a_2) + A(\varepsilon + rh_0).$$

This in turn (instead of (4.40)) leads to the estimate

$$h(x_1, x_2) \leq Ah(a_1, a_2) + 3Arh_0,$$

and this proves the assertion (2.17). In order to prove (2.18) we proceed in precisely the same way as we did it for proving (2.23) in Theorem 2.5. The only difference is that instead of (4.42) we have

$$h\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) \leq 2h(z_1, z_2) \leq 2\varepsilon$$

which using now (2.16) leads to the same contradiction (4.43). Thus, (2.18) follows. □

Chapter 5

Proof of the results from Sections 2.3 and 2.4

5.1 Heights

By the Product formula we have for any number field K and any $x \in K^*$ that

$$h(x) = \sum_{v \in M_K} \max(0, \log |x|_v) = \frac{1}{2} \sum_{v \in M_K} |\log |x|_v|. \tag{5.1}$$

Recall that we have defined

$$h(\mathbf{x}) := \sum_{i=1}^{n} h(x_i)$$

for $\mathbf{x} = (x_1, \dots, x_N) \in (\overline{\mathbb{Q}}^*)^N$. Further, for $\xi \in \mathbb{Q}$ we define $\mathbf{x}^\xi := (x_1^\xi, \dots, x_N^\xi)$. The point $\mathbf{x}^\xi$ is determined only up to multiplication with $(\overline{\mathbb{Q}}^*_{\text{tors}})^N$ where $\overline{\mathbb{Q}}^*_{\text{tors}} = \{\boldsymbol{\rho} \in \overline{\mathbb{Q}}^* \ : \ \exists m \in \mathbb{Z}_{>0} \ \text{ with } \ \boldsymbol{\rho}^m = 1\}$. But $h(\mathbf{x}^\xi)$ is well defined. It now follows easily that

$$h(\mathbf{xy}) \le h(\mathbf{x}) + h(\mathbf{y}), \ h(\mathbf{x}^\xi) = |\xi| h(\mathbf{x}) \text{ for } \mathbf{x}, \mathbf{y} \in (\overline{\mathbb{Q}}^*)^N, \xi \in \mathbb{Q},$$

and $h(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in (\overline{\mathbb{Q}}^*_{\text{tors}})^N$.

We define several heights for polynomials. Let f be a non-zero polynomial with coefficients in $\overline{\mathbb{Q}}$, and let $a_1, \dots, a_R$ be its non-zero coefficients. Choose a number field K such that $a_1, \dots, a_R \in K$. Recall that for every infinite place v of K there is an embedding $\sigma_v : K \hookrightarrow \mathbb{C}$ such that $|\cdot|_v = |\sigma_v(\cdot)|^{\varepsilon_v}$, where $\varepsilon_v := [K_v : \mathbb{R}]/[K : \mathbb{Q}]$. For $v \in M_K$ we put $\|f\|_v := \max_{1 \le i \le R} |a_i|_v$. Further, for every infinite place v of K and every $l \ge 1$ we put $\|f\|_{v,l} := \left(\sum_{i=1}^{R} |\sigma_v(a_i)|^l \right)^{\varepsilon_v / l}$. We have already defined

$$h(f) := \sum_{v \in M_K} \log \|f\|_v.$$

In addition, we define the heights

$$h_l(f) := \sum_{v|\infty} \log \|f\|_{v,l} + \sum_{v\nmid\infty} \log \|f\|_v \text{ for } l \geq 1,$$

and the Gauss-Mahler height

$$h_{GM}(f) := \sum_{v|\infty} \varepsilon_v \log M(f^{\sigma_v}) + \sum_{v\nmid\infty} \log \|f\|_v,$$

where f^σ is the polynomial obtained by applying σ to the coefficients of f and $M(\cdot)$ denotes the Mahler measure of a polynomial with complex coefficients. None of these heights depends on the choice of K. We have

$$h_{GM}(f) \leq h_1(f), \quad h(f) \leq h_1(f) \leq h(f) + \log R, \tag{5.2}$$

where R is the number of non-zero coefficients of f. Further, for any non-zero polynomial $P \in \overline{\mathbb{Q}}[X]$ and any root ζ of P we have

$$h(\zeta) \leq h_{GM}(P) \leq h_1(P). \tag{5.3}$$

We use also exponential heights $H(x) = \exp(h(x))$ for $x \in \overline{\mathbb{Q}}$, and likewise $H(f), H_l(f)$, $H_{GM}(f)$ for polynomials f with coefficients in $\overline{\mathbb{Q}}$.

5.2 Main tools

In this section we have collected the tools needed in the sequel.

We start with some results from [8] that have been derived from lower bounds for linear forms in logarithms. Let K be an algebraic number field of degree d, M_K the set of places on K, and G a finitely generated multiplicative subgroup of K^* of rank $t > 0$. We fix a set of (not necessarily multiplicatively independent) generators $\{\xi_1, \ldots, \xi_r\}$ of G modulo G_{tors} and put

$$Q := \prod_{i=1}^{r} \max\big(1, h(\xi_i)\big). \tag{5.4}$$

Let $P(v)$ ($v \in M_K$) be given by A.2.0b, i.e., $P(v) := 2$ if v is infinite and $P(v) := \#\mathcal{O}_K/\mathfrak{p}_v$ if v is finite, where $\mathfrak{p}_v$ is the prime ideal of $\mathcal{O}_K$ corresponding to v.

Lastly, let

$$c(r,d) := 20(16ed)^{3(r+2)} \left(\frac{r}{e}\right)^r.$$

Lemma 5.1. *Let $\alpha \in K^*$ with $\max(h(\alpha), 1) \leq H$, $v \in M_K$, and $0 < \kappa \leq 1$. Then for every $\xi \in G$ with $\alpha\xi \neq 1$ and*

$$\log |1 - \alpha\xi|_v < -\kappa h(\xi) \tag{5.5}$$

we have $h(\xi) < C_4(\kappa) \cdot H$, where

$$\begin{aligned} C_4(\kappa) \;&:=\; (c(r,d)/\kappa)\frac{P(v)}{\log P(v)}Q \cdot \\ &\qquad \cdot \max\{\log(c(r,d)P(v)/\kappa), \log^* Q\}. \end{aligned}$$

Proof. This is our Theorem 2.7, with instead of $c(r,d)$ a constant c depending also on the rank t of G. However, using $t \leq r$ an easy computation proves the estimate of our lemma. □

We keep the notation from above. In addition, let S be a finite set of places of K containing all infinite places such that $G \subset \mathcal{O}_S^*$. Put $s := \#S$ and define $\mathbf{P}$ by (2.5), that is $\mathbf{P} := \max_{v \in S} P(v)$. Consider the equation

$$\alpha x + \beta y = 1 \;\text{ in } x \in G,\, y \in \mathcal{O}_S^*, \tag{5.6}$$

where $\alpha, \beta \in K^*$ with $\max(h(\alpha), h(\beta), 1) \leq H$.

Lemma 5.2. *For every solution $x \in G, y \in \mathcal{O}_S^*$ of (5.6) we have*

$$\max\left(h(x), h(y)\right) < C_5 H, \tag{5.7}$$

where

$$C_5 := c(r,d) \cdot \frac{s\mathbf{P}}{\log \mathbf{P}} Q \cdot \max\{\log(c(r,d)s\mathbf{P}), \log^* Q\}.$$

Proof. This is our Theorem 2.2, again with a constant c depending on the rank t of G which we bounded above using $t \leq r$. □

Below we have collected some results on heights of algebraic points.

Lemma 5.3. *(i) Let $\alpha, \beta \in \overline{\mathbb{Q}}^*$. Then there are at most two points $\mathbf{x} = (x, y) \in (\overline{\mathbb{Q}}^*)^2$ such that*

$$\alpha x + \beta y = 1, \quad h(\mathbf{x}) \leq 0.03.$$

(ii) Let $f(X, Y) \in \overline{\mathbb{Q}}[X, Y]$ be an irreducible polynomial which is not a binomial. Then the number of points $\mathbf{x} = (x, y) \in (\overline{\mathbb{Q}}^*)^2$ with

$$f(x, y) = 0, \quad h(\mathbf{x}) \leq \left(2^{47} \deg_s f (\log \deg_s f)^5\right)^{-1}$$

is at most

$$2^{50} \deg_s f (\log \deg_s f)^6.$$

Proof. (i) Beukers and Zagier [14, Corollary 2.4] proved that if there are three points $(x_1, y_1), (x_2, y_2), (x_3, y_3) \in (\overline{\mathbb{Q}}^*)^2$ satisfying $\alpha x_i + \beta y_i = 1$ for $i = 1, 2, 3$, then $\sum_{i=1}^{3} h(x_i, y_i) \geq \log \rho$, where ρ denotes the real root of $\rho^{-6} + \frac{1}{2}\rho^{-2} = 1$ which is larger than 1. We have $\log \rho > 0.09$.
(ii) This is proved by Pontreau in [59, Proposition 5.1] (see also [58, Proposition 3.3]). □

Our last height result is an effective version of a special case of Bézout's Theorem.

Lemma 5.4. *Let $f, g \in \overline{\mathbb{Q}}[X, Y]$ be two coprime polynomials. Then for every common zero $\mathbf{x} = (x, y)$ of f and g we have*

$$h(\mathbf{x}) \leq \deg_s g \cdot h_{GM}(f) + \deg_s f \cdot h_1(g).$$

Proof. See [58, Lemma 3.7]. □

5.3 Proof of Theorem 2.8

We follow the proof of Bombieri and Gubler [18, Thm. 5.4.5, pp. 147–148].

We denote the partial degrees of f with respect to X, Y by δ_X, δ_Y, respectively, and put $\delta := \deg_s f = \delta_X + \delta_Y$. From our assumptions it follows that f is irreducible over $\overline{\mathbb{Q}}$, that f has at least three non-zero terms, and hence that $\delta_X \geq 1$, $\delta_Y \geq 1$.

We assume that one of the coefficients of f is 1 which is no loss of generality since the height of a polynomial is invariant under multiplication by a scalar.

Recall that we allow that f has its coefficients in $\overline{\mathbb{Q}}$; this will be needed in the proofs of Theorems 2.9, 2.10. But in fact there is no loss of generality to assume that $f \in K[X, Y]$. To see this, suppose that $f \notin K[X, Y]$. Then there is $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ such that the polynomial f^σ obtained by applying σ to the coefficients of f is distinct from f. Since one of the coefficients of f is 1, f^σ cannot be proportional to f, and since f is irreducible over $\overline{\mathbb{Q}}$, f^σ has to be coprime to f. Now if $\mathbf{x} \in \Gamma$ is a zero of f then it is also a zero of f^σ. Thus, by Lemma 5.4, (5.2), noting that $\deg_s f = \deg_s f^\sigma = \delta$, it follows that

$$h(\mathbf{x}) \leq \delta(h_{GM}(f^\sigma) + h_1(f)) \leq 2\delta(H + 2\log \delta)$$

and this is much sharper than the bound from Theorem 2.8.

Write

$$f(X, Y) = \sum_{(i,j) \in \mathcal{F}} a_{ij} X^i Y^j \text{ with } a_{ij} \in K^* \text{ for } (i, j) \in \mathcal{F}, \tag{5.8}$$

where $\mathcal{F}$ is a subset of $\{0,\dots,\delta_X\}\times\{0,\dots,\delta_Y\}$. Thus,

$$\#\mathcal{F}\le(\delta_X+1)(\delta_Y+1)\le\delta^2.$$

The height $H(f)$ remains unaltered under multiplication by a_{ij}^{-1} for any $(i,j)\in\mathcal{F}$, so we have for any place $v\in M_K$ and any two pairs $(i,j),(p,q)\in\mathcal{F}$,

$$|a_{pq}/a_{ij}|_v\le\max_{k,l}|a_{kl}/a_{ij}|_v\le H(f)$$

and by interchanging the role of a_{pq},a_{ij},

$$H(f)^{-1}\le|a_{pq}/a_{ij}^{-1}|_v\le H(f). \tag{5.9}$$

Put $s:=\#S$. Take a point $\mathbf{x}=(x,y)\in\mathcal{C}\cap\Gamma$ with

$$H(\mathbf{x})\ge(\delta^2H(f))^{24s\delta^4}. \tag{5.10}$$

Notice that the logarithm of the right-hand side is much smaller than the upper bound C_1H from our Theorem. By the product formula we have

$$\begin{aligned}H(\mathbf{x})^2=(H(x)H(y))^2 &= \prod_{v\in S}\max(|x|_v,|x|_v^{-1})\max(|y|_v,|y|_v^{-1})\\ &\le \prod_{v\in S}\max(|x|_v,|x|_v^{-1},|y|_v,|y|_v^{-1})^2.\end{aligned}$$

Thus, there exists $v\in S$ such that

$$\max(|x|_v,|x|_v^{-1},|y|_v,|y|_v^{-1})\ge H(\mathbf{x})^{1/s}\ge(\delta^2H(f))^{24\delta^4}.$$

Replacing x by $x^{\pm1},y^{\pm1}$ and correspondingly f by a polynomial $\tilde f$ with $\tilde f(x^{\pm1},y^{\pm1})=0$ (which has the same partial degrees and height as f), we see that there is no loss of generality to assume that $\min(|x|_v,|y|_v)\ge1$ and moreover,

$$\max(|x|_v,|y|_v)\ge H(\mathbf{x})^{1/s}\ge(\delta^2H(f))^{24\delta^4}. \tag{5.11}$$

Now let us order the pairs in $\mathcal{F}$ according to

$$|x^py^q|_v\ge|x^ry^s|_v\ge\cdots.$$

Recall that f is not a binomial. Hence $\mathcal{F}$ contains pairs other than (p,q), (r,s). Further, $\delta_X,\delta_Y\ge1$ so $\mathcal{F}$ contains pairs (i,j) with $i>0$ and pairs with $j>0$. Using also $\min(|x|_v,|y|_v)\ge1$, it follows that $|x^py^q|_v\ge\max(|x|_v,|y|_v)$. Now (5.11) gives

$$|x^py^q|_v\ge H(\mathbf{x})^{\frac1s}\ge(\delta^2H(f))^{24\delta^4}. \tag{5.12}$$

We compare $|x^p y^q|_v$, $|x^r y^s|_v$. Using that $f(x,y)=0$ and also (5.8), (5.9), and the fact that $\#\mathcal{F} \le \delta^2$, we obtain

$$|x^p y^q|_v \le \delta^2 \max_{(i,j)\in\mathcal{F}} |a_{ij}|_v |a_{pq}|_v^{-1} |x^i y^j|_v \le \delta^2 H(f) |x^r y^s|_v.$$

Hence

$$1 \le \left|x^{p-r} y^{q-s}\right|_v \le \delta^2 H(f). \tag{5.13}$$

We claim that (p,q) and (r,s) are linearly independent. Indeed, assume there exists $u \in \mathbb{Q}\setminus\{1\}$ such that $(up,uq)=(r,s)$. We deduce from (5.13)

$$|x^p y^q|_v^{|1-u|} \le \delta^2 H(f).$$

We note that from $p,q \le \delta-1$ it follows $|u-1| \ge \frac{1}{\delta-1}$, thus

$$|x^p y^q|_v \le (\delta^2 H(f))^{\delta-1}$$

which contradicts (5.12).

Hence for all $(i,j)\in\mathcal{F}$ there are $A_{ij}, B_{ij} \in \mathbb{Q}$ with

$$i = A_{ij}p + B_{ij}r, \quad j = A_{ij}q + B_{ij}s.$$

Let $(i,j)\in\mathcal{F}$. Then using

$$x^i y^j = (x^p y^q)^{A_{ij}+B_{ij}} (x^{r-p} y^{s-q})^{B_{ij}} \tag{5.14}$$

and (5.13), we get

$$\begin{aligned} |x^p y^q|_v &\ge |x^i y^j|_v = |x^p y^q|_v^{A_{ij}+B_{ij}} |x^{r-p} y^{s-q}|_v^{B_{ij}} \\ &\ge |x^p y^q|_v^{A_{ij}+B_{ij}} \cdot (\delta^2 H(f))^{-|B_{ij}|}. \end{aligned}$$

Put $D = |ps-qr|$. Then $D, D\cdot A_{ij} = is - jr$ and $D\cdot B_{ij} = pj - qi \in \mathbb{Z}$ and moreover, $|D| \le (\delta-1)^2$, $|DA_{ij}| \le (\delta-1)^2$, $|DB_{ij}| \le (\delta-1)^2$. Therefore,

$$|x^p y^q|_v^{D-D(A_{ij}+B_{ij})} \ge (\delta^2 H(f))^{-(\delta-1)^2}.$$

Since $|x^p y^q|_v > (\delta^2 H(f))^{(\delta-1)^2}$ (by (5.12)) the integer $D - D(A_{ij}+B_{ij})$ is non-negative, in other words $A_{ij}+B_{ij}=1$ or $A_{ij}+B_{ij} \le 1-\frac{1}{D}$. Now define $\mathcal{I}$ to be the set of $(i,j)\in\mathcal{F}$ such that $A_{ij}+B_{ij}=1$. The set $\mathcal{I}$ contains at least the pairs (p,q) and (r,s). Choose a D-th root $z^{1/D}$ of $z := x^{r-p}y^{s-q}$. Then by (5.14) we have

$$\begin{aligned} &0 = f(x,y) = x^p y^q R(z^{1/D}) + Q(x,y) \\ &\text{with } R(Z) := \sum_{(i,j)\in\mathcal{I}} a_{ij} Z^{DB_{ij}}, \; Q(X,Y) := \sum_{(i,j)\in\mathcal{F}\setminus\mathcal{I}} a_{ij} X^i Y^j. \end{aligned} \tag{5.15}$$

Let $m := -\min\{DB_{ij} : (i,j) \in \mathcal{I}\}$ and put $R^*(Z) := Z^m R(Z)$. Thus $R^*(Z)$ is a polynomial with $R^*(0) \neq 0$. Since $\mathcal{I}$ contains at least two pairs, the polynomial R^* is non-constant. Choose an extension of $|\cdot|_v$ to $\overline{\mathbb{Q}}$. We proceed to estimate from above $|R^*(z^{1/D})|_v$.

Let $(i,j) \in \mathcal{F} \setminus \mathcal{I}$. Then by (5.14), $A_{ij} + B_{ij} \leq 1 - \frac{1}{D}$, $|B_{ij}| \leq (\delta-1)^2/D$, (5.13) we have

$$\begin{aligned} |x^i y^j|_v &= |x^p y^q|_v^{A_{ij}+B_{ij}} \cdot \left|x^{r-p} y^{q-s}\right|_v^{B_{ij}} \\ &\leq |x^p y^q|_v^{1-\frac{1}{D}} \cdot (\delta^2 H(f))^{(\delta-1)^2/D}. \end{aligned}$$

Hence

$$|Q(x,y)|_v \leq |x^p y^q|_v^{1-1/D} \cdot (\delta^2 H(f))^{1+(\delta-1)^2/D}.$$

Using this estimate together with (5.13), (5.12), we obtain

$$\begin{aligned} |R^*(z^{1/D})|_v &= |z|_v^{m/D} |R(z^{1/D})|_v = |z|_v^{m/D} |Q(x,y)|_v \\ &\leq (\delta^2 H(f))^{\delta^2/D} |x^p y^q|_v^{-1/D} (\delta^2 H(f))^{1+(\delta-1)^2/D} \\ &\leq (\delta^2 H(f))^{(3\delta^2)/D} H(\mathbf{x})^{-1/sD}. \end{aligned}$$

It is useful to observe here that in the above argument the D-th root $z^{1/D}$ was chosen arbitrarily. Thus, we have

$$|\prod_{\rho} R^*(\rho z^{1/D})|_v \leq (\delta^2 H(f))^{3\delta^2} H(\mathbf{x})^{-1/s} \tag{5.16}$$

where the product is taken over all D-th roots of unity.

Notice that the constant term of R^* is a coefficient of f, say a_{i_0,j_0}. By dividing f by a_{i_0,j_0} as we may since it does not affect the above estimates, we get that the constant term of R^* is 1. Thus we have

$$R^*(Z) = \prod_{\zeta} (1 - \zeta^{-1} Z)$$

where the product is taken over all zeros of R^*. So

$$\prod_{\rho} R^*(\rho z^{1/D}) = \prod_{\zeta} (1 - \zeta^{-D} z).$$

Choose some ζ for which $|1 - \zeta^{-D} z|_v$ is minimal. Using (5.16), (5.12), and also that R^* has degree at most $2\delta^2$ and that $H(z) \leq H(\mathbf{x})^\delta$ we arrive at

$$\begin{aligned} |1 - \zeta^{-D} z|_v &\leq \left\{(\delta^2 H(f))^{3\delta^2} H(\mathbf{x})^{-1/s}\right\}^{1/\deg R^*} \\ &\leq \left(H(\mathbf{x})^{-2/3s}\right)^{1/2\delta^2} \leq H(z)^{-1/3s\delta^3}. \end{aligned}$$

The number ζ^{-D} may lie outside K. Let $K' = K(\zeta^D)$. Then $[K' : K] \le 2\delta^2$ and there is a place v' of K' lying above v such that $|\gamma|_{v'} = |\gamma|_v^{[K'_{v'}:K_v]/[K':K]}$ for $\gamma \in K'$ where $|\cdot|_{v'}$ is normalized with respect to K'. Thus we finally obtain

$$\log|1 - \zeta^{-D}z|_{v'} \le -\frac{1}{6s\delta^5} \cdot h(z). \tag{5.17}$$

Now we apply Lemma 5.1 to (5.17) with $\alpha = \zeta^{-D}$, $\kappa = (6s\delta^5)^{-1}$, K' instead of K, v' instead of v and we take for G the group $\{x^{r-p}y^{s-q} : (x, y) \in \Gamma\}$. Notice that by (5.3), (5.2),

$$\begin{aligned} &h(\zeta^D) \le Dh_1(R^*) \le \delta^2 h_1(f) \le \delta^2(H + 2\log\delta), \\ &[K' : \mathbb{Q}] \le 2\delta^2 d,\ P(v') \le P(v)^{2\delta^2}. \end{aligned}$$

So in the bound $C_4(\kappa)H$ from Lemma 5.1 we have to replace H by $\delta^2(H + 2\log\delta)$, κ by $(6\delta^5 s)^{-1}$, d by $2\delta^2 d$ and $P(v)$ by $P(v') \le P(v)^{2\delta^2} \le \mathbf{P}^{2\delta^2}$. Further, if $\{\mathbf{w}_i = (w_{1i}, w_{2i}) : i = 1, \dots, r\}$ is a basis of Γ modulo Γ_{tors}, the group G is generated modulo G_{tors} by the numbers $\xi_i := w_{1i}^{r-p} w_{2i}^{s-q}$ $(i = 1, \dots, r)$ and so for the quantity Q defined by (5.4) we have

$$Q = \prod_{i=1}^{r} \max(1, h(\xi_i)) \le (\delta h_0)^r.$$

A straightforward computation shows that with these replacements for H, κ, $P(v)$ and the upper bound for Q, the constant $c(r, d)$ becomes $c' := 20(32e\delta^2 d)^{3r+6}(32^3 e^2 r)^r$, and $C_4(\kappa)$ can be estimated from above by

$$\begin{aligned} c' \cdot 6\delta^5 s \cdot \frac{\mathbf{P}^{2\delta^2}}{2\delta^2 \log \mathbf{P}} \cdot (\delta h_0)^r \cdot \\ \cdot \max\Big(\log(c'\mathbf{P}^{2\delta^2} \cdot 6\delta^5 s), \log^*\big((\delta h_0)^r\big)\Big). \end{aligned}$$

Using that the maximum is at most $100r\delta^2 \log^*\big(\max(\delta ds\mathbf{P}, \delta h_0)\big)$, we obtain for $C_4(\kappa)$ the upper bound

$$C := e^{36}(e^{13}r)^r \delta^{7r+17} d^{3r+6} s h_0^r \cdot \frac{\mathbf{P}^{2\delta^2}}{\log \mathbf{P}} \cdot r^2 \log^*\big(\max(\delta ds\mathbf{P}, \delta h_0)\big).$$

Thus, if $z \ne \zeta^D$ we get

$$h(z) < C \max(1, h(\zeta^{-D})) \le C\delta^2(H + 2\log\delta),$$

while if $z = \zeta^D$ we get $h(z) \le \delta^2(H + 2\log\delta)$ which is much smaller.

We proved that $\mathbf{x} = (x, y)$ verifies an equation $x^{r-p}y^{s-q} = \mu$ for some $\mu \in K$ with

$$h(\mu) \leq C \cdot \delta^2(H + 2\log\delta).$$

Since f is irreducible and not a binomial, we can apply Lemma 5.4 and obtain, using $h_{GM}(X^rY^s - \mu X^pY^q) = h(\mu)$, $h_1(f) \leq H + 2\log\delta$, the upper bound

$$\begin{aligned} h(\mathbf{x}) &\leq \delta(h_1(f) + h(\mu)) \leq \delta(\delta^2 C + 1) \cdot (H + 2\log\delta) \\ &\leq 3\delta^4 CH \leq C_1 H. \end{aligned}$$

Our Theorem follows.

5.4 Proof of Theorems 2.9 and 2.10

Theorems 2.9 and 2.10 are proved in the same manner. We prove only Theorem 2.10 and then indicate the necessary modifications to obtain a proof of Theorem 2.9.

Proof of Theorem 2.10. Let $\mathbf{x} \in \mathcal{C} \cap C(\overline{\Gamma}, \varepsilon)$ with the value of ε given by (2.29). We may write $\mathbf{x} = \mathbf{yz}$ with $\mathbf{y} \in \overline{\Gamma}$ and $\mathbf{z} \in (\overline{\mathbb{Q}}^*)^2$ with $h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y}))$. We may further split up $\mathbf{y}$ as

$$\mathbf{y} = \mathbf{vw} \quad \text{with } \mathbf{v} \in \Gamma, \mathbf{w} = \prod_{i=1}^{r} \mathbf{w}_i^{\gamma_i}, \tag{5.18}$$

where $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \leq \frac{1}{2}$. Here $\mathbf{w}$ is determined only up to a factor from $(\overline{\mathbb{Q}}^*)^2_{\text{tors}}$ but this will not cause problems.

Define now a new polynomial $f^*(\mathbf{V}) := f(\mathbf{wz} \cdot \mathbf{V})$. Notice that $f^*(\mathbf{v}) = 0$. First observe that $\deg_s f^* = \deg_s f$ which we write again as δ. Further, $h(f^*) \leq h(f) + \delta h(\mathbf{wz}) \leq h(f) + \delta(h(\mathbf{w}) + h(\mathbf{z}))$. By applying Theorem 2.8 to f^* we obtain

$$\begin{aligned} h(\mathbf{v}) &\leq C_1 H + C_1\delta(h(\mathbf{w}) + h(\mathbf{z})) \\ &\leq C_1 H + C_1\delta \cdot \Big(\varepsilon\big(1 + h(\mathbf{vw})\big) + h(\mathbf{w})\Big) \\ &\leq C_1\delta\varepsilon h(\mathbf{v}) + C_1\delta(\varepsilon + (1+\varepsilon)h(\mathbf{w})) + C_1 H. \end{aligned} \tag{5.19}$$

Here it is essential that the bound of Theorem 2.8 does not depend on the field generated by the coefficients of f^*. Further,

$$\begin{aligned} h(\mathbf{x}) &\leq h(\mathbf{vw}) + \varepsilon \cdot (1 + h(\mathbf{vw})) \\ &\leq \varepsilon + (1+\varepsilon) \cdot (h(\mathbf{v}) + h(\mathbf{w})) \\ &\leq \varepsilon + (1+\varepsilon)h(\mathbf{w}) + (1+\varepsilon)h(\mathbf{v}). \end{aligned} \tag{5.20}$$

By our choice of ε we have $(1+\varepsilon)(1-C_1\varepsilon\delta)^{-1} \le 2$. Further,

$$h(\mathbf{w}) \le \sum_{i=1}^{r} |\gamma_i| \cdot h(\mathbf{w}_i) \le \frac{1}{2} r h_0.$$

By inserting this bound as well as the upper bound for $h(\mathbf{v})$ resulting from (5.19) into (5.20), we obtain

$$\begin{aligned} h(\mathbf{x}) &\le \Big(\varepsilon + (1+\varepsilon)h(\mathbf{w})\Big) \cdot \Big(1 + 2C_1\delta\Big) + 2C_1 H \\ &\le \Big(\varepsilon + (1+\varepsilon) r h_0/2\Big) \cdot \Big(1 + 2C_1\delta\Big) + 2C_1 H \\ &\le 2 r h_0 \delta C_1 + 2C_1 H. \end{aligned} \tag{5.21}$$

This is the upper bound for $h(\mathbf{x})$ in Theorem 2.10.

We now estimate from above $[L(\mathbf{x}) : L]$ where L is the number field generated by Γ and the coefficients of f. This degree is equal to the number of distinct points among $\sigma(\mathbf{x})$ where $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. So we have to estimate from above the latter. $\mathbf{y}$, $\mathbf{v}$, $\mathbf{w}$ will be as above.

Pick $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Define $g(\mathbf{X}) := f(\mathbf{x} \cdot \mathbf{X})$. Notice that $\deg_s g = \deg_s f = \delta$. Since some integer power of $\mathbf{y}$ belongs to $\Gamma \subseteq L^2$ and σ is a L-isomorphism, we infer that $\sigma(\mathbf{y})\mathbf{y}^{-1}$ is a root of unity. It follows that

$$h(\sigma(\mathbf{x})\mathbf{x}^{-1}) = h(\sigma(\mathbf{z})\mathbf{z}^{-1}) \le 2h(\mathbf{z}).$$

The point $\sigma(\mathbf{x})\mathbf{x}^{-1}$ belongs to the curve defined by g. So, under the assumption

$$2h(\mathbf{z}) \le \Big(2^{47}\delta(\log\delta)^5\Big)^{-1} \tag{5.22}$$

we deduce from Lemma 5.3 (ii) that the number of distinct points $\sigma(\mathbf{x})$ is at most

$$2^{50}\delta^2(\log\delta)^6$$

and this is precisely the upper bound from Theorem 2.10.

It remains to prove (5.22). We have $h(\mathbf{z}) \le \varepsilon \cdot \Big(1 + h(\mathbf{w}) + h(\mathbf{v})\Big)$ so as in (5.19) we obtain

$$h(\mathbf{z}) \le \varepsilon \cdot \Big(1 + h(\mathbf{w}) + C_1 H + C_1\delta \cdot h(\mathbf{w}) + h(\mathbf{z})\Big)$$

implying

$$\Big(1 - \varepsilon C_1\delta\Big)h(\mathbf{z}) \le \varepsilon \cdot \Big(\big(1 + C_1\delta\big) \cdot h(\mathbf{w}) + 1 + C_1 H\Big).$$

Then inserting $h(\mathbf{w}) \le \frac{1}{2} r h_0$ and using (2.29) we get

$$h(\mathbf{z}) \le \varepsilon \cdot \big(C_1\delta r h_0 + 2C_1 H\big). \tag{5.23}$$

Now our choice of ε in (2.29) implies indeed (5.22). □

Proof of Theorem 2.9. The proof is very similar to that of Theorem 2.10. We indicate only the necessary changes.

So let $\mathbf{x} \in \mathcal{C}(\overline{\mathbb{Q}}) \cap \overline{\Gamma}_\varepsilon$ with ε given by (2.28). Then $\mathbf{x} = \mathbf{yz}$ with $\mathbf{y} \in \overline{\Gamma}_\varepsilon$ and $h(\mathbf{z}) < \varepsilon$. Write again $\mathbf{y} = \mathbf{vw}$ with $\mathbf{v} \in \Gamma$ and $\mathbf{w} = \prod_{i=1}^r \mathbf{w}_i^{\gamma_i}$ with $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \leq \frac{1}{2}$.

Now using $h(\mathbf{z}) < \varepsilon$ we obtain instead of (5.19),

$$h(\mathbf{v}) \leq C_1\delta(h(\mathbf{w}) + \varepsilon) + C_1 H.$$

Further,

$$h(\mathbf{x}) \leq h(\mathbf{v}) + h(\mathbf{w}) + h(\mathbf{z}) \leq (1 + \delta C_1)h(\mathbf{w}) + \varepsilon + C_1 H$$

and by inserting $h(\mathbf{w}) \leq \frac{1}{2} r h_0$, we obtain

$$h(\mathbf{x}) \leq r h_0 \delta C_1 + C_1 H$$

which is the bound from Theorem 2.9.

We now estimate from above $[L(\mathbf{x}) : L]$ and for this we have to estimate the number of distinct points among $\sigma(\mathbf{x})$, $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. As above we have

$$h(\sigma(\mathbf{x})\mathbf{x}^{-1}) = h(\sigma(\mathbf{z})\mathbf{z}^{-1}) < 2\varepsilon.$$

Thanks to our choice of ε in (2.28) we have (5.22), and our upper bound for $[L(\mathbf{x}) : L]$ follows in the same manner as above. □

5.5 Points in translates of algebraic groups

In the present section we prove effective results on the intersection of Γ or $\overline{\Gamma}_\varepsilon$ with a translate $\mathbf{x}_0\mathcal{H}$, where Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$, $\varepsilon > 0$, $\mathbf{x}_0 \in (\overline{\mathbb{Q}}^*)^N$ is fixed and $\mathcal{H}$ is a proper algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$. In fact we show that if $\mathbf{x}_0\mathcal{H}$ contains a point from Γ or $\overline{\Gamma}_\varepsilon$ then it contains such a point with height and degree below some effectively computable constants. Thus, it can be decided effectively whether or not $\mathbf{x}_0\mathcal{H}$ contains points from Γ or $\overline{\Gamma}_\varepsilon$.

For $\mathbf{x} = (x_1, \ldots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ and an $N \times M$-matrix $A = \left(a_{ij}\right)_{1 \leq i \leq M, 1 \leq j \leq N}$, with $a_{ij} \in \mathbb{Z}$ we define $\mathbf{x}^A \in (\overline{\mathbb{Q}}^*)^M$ by

$$\mathbf{x}^A := \left(x_1^{a_{11}} \ldots x_N^{a_{N1}}, \ldots, x_1^{a_{1M}} \ldots x_N^{a_{NM}}\right).$$

Thus, $(\mathbf{x}^A)^B = \mathbf{x}^{AB}$ whenever the product of the matrices A, B is defined. It is well-known that for every $(N - M)$-dimensional algebraic subgroup $\mathcal{H}$ of $(\overline{\mathbb{Q}}^*)^N$ there is an

integer $N \times M$-matrix A of rank M such that $\mathcal{H}$ is the set of points $\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N$ with $\mathbf{x}^A = \mathbf{1} = (1, \ldots, 1)$ (M times) (see for instance [18, Theorem 3.2.19]). Moreover, every translate of $\mathcal{H}$ can be described as the set of solutions of $\mathbf{x}^A = \mathbf{c}$ for some fixed $\mathbf{c} \in (\overline{\mathbb{Q}}^*)^M$. (See for instance again [18].)

As before, we choose a basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ of Γ modulo Γ_{tors}. Let K be the smallest number field such that $\Gamma \subset (K^*)^N$ and let S be the smallest set of places of K that contains all infinite places and such that $\Gamma \subset (\mathcal{O}_S^*)^N$. Put

$$h_0 := \max\{1, h(\mathbf{w}_1), \ldots, h(\mathbf{w}_r)\},\ d := [K : \mathbb{Q}],\ s := \#S.$$

Notice that by the product formula we have for $\mathbf{x} = (x_1, \ldots, x_N) \in \Gamma$,

$$h(\mathbf{x}) = \frac{1}{2} \sum_{v \in S} \sum_{i=1}^{N} |\log |x_i|_v|. \tag{5.24}$$

Let $A = \big(a_{ij}\big)_{1 \le i \le N,\, 1 \le j \le M}$ be an integer $N \times M$-matrix, where we do not require that A has rank M. Further, let $\mathbf{c}$ be a fixed point of $(\overline{\mathbb{Q}}^*)^M$, and δ, H reals such that

$$\max_{i,j} |a_{ij}| \le \delta, \quad \max(1, h(\mathbf{c})) \le H.$$

Let $c(d)$ be the constant from Lemma 4.1.

Our first result is as follows.

Proposition 5.5. *Assume that*

$$\mathbf{x}^A = \mathbf{c} \quad \textit{in } \mathbf{x} \in \Gamma \tag{5.25}$$

is solvable. Then (5.25) has a solution $\mathbf{x}_0 \in \Gamma$ such that

$$h(\mathbf{x}_0) \le h_0 \cdot \big(r 4^r c(d) M \delta h_0\big)^r \cdot H.$$

In the proof we need some results on lattice points. We start with recalling a result of Schlickewei [66, Proposition 4.2].

Lemma 5.6. *Let Λ be a discrete subgroup of rank r in $\mathbb{R}^m$ and $\|\cdot\|$ a norm on $\mathbb{R}^m$. Then there exists a basis $\mathbf{a}_1, \ldots, \mathbf{a}_r$ of Λ such that for any $x_1, \ldots, x_r \in \mathbb{Z}$ we have*

$$\|x_1\mathbf{a}_1 + \cdots + x_r\mathbf{a}_r\| \ge 4^{-r} \max\{|x_1|\|\mathbf{a}_1\|, \ldots, |x_r|\|\mathbf{a}_r\|\}. \tag{5.26}$$

Proof. Schlickewei proved this only for $\mathbb{Z}^r$ instead of arbitrary lattices Λ, but using a suitable linear transformation the above more general result follows in a straightforward way. □

In the sequel let $\|\cdot\|_l$ denote the usual l-norm defined by $\|\mathbf{x}\|_l = \left(\sum_i |x_i|^l\right)^{1/l}$ if $1 \le l < \infty$ and $\|\mathbf{x}\|_\infty = \max_i |x_i|$.

Lemma 5.7. *Let U be an $r \times k$ integer matrix of rank k and $\mathbf{m} \in \mathbb{Z}^k$. Further, let R, V be reals such that the coordinates of $\mathbf{m}$ have absolute values at most R and the entries of U have absolute values at most V. Suppose that the equation*

$$\mathbf{x}U = \mathbf{m} \quad \textit{in } \mathbf{x} \in \mathbb{Z}^r \tag{5.27}$$

has a solution. Then equation (5.27) has a solution $\mathbf{x}_0 \in \mathbb{Z}^r$ such that

$$\|\mathbf{x}_0\|_\infty \le k^{k/2} V^{k-1} \max(V, R).$$

Proof. According to a result of Borosh, Flahive, Rubin and Treybig [20], (5.27) has a solution $\mathbf{x}_0$ with $\|\mathbf{x}_0\|_\infty \le W$, where W is the maximum of the absolute values of the minors of the augmented matrix with U on the first r rows and $\mathbf{m}$ on the last row. Now our Lemma follows easily by applying Hadamard's inequality. □

Proof of Proposition 5.5. Put $s := \#S$. For any positive integer t, we denote by φ_t the group homomorphism from $(\mathcal{O}_S^*)^t$ to $\mathbb{R}^{st}$, given by

$$\varphi_t : \ \mathbf{x} \mapsto (\log |x_i|_v : v \in S, i = 1, \dots, t),$$

where we have written $\mathbf{x} = (x_1, \dots, x_t)$. Further, denote by $\|\cdot\|$ the 1-norm on $\mathbb{R}^{Ns}$ and by $\|\cdot\|^*$ the 1-norm on $\mathbb{R}^{Ms}$.

The kernel of $\varphi := \varphi_N|_\Gamma$ is Γ_{tors}, and the image Λ of φ in $\mathbb{R}^{Ns}$ is a discrete subgroup of rank r. Equation (5.25) can be written in the form

$$\mathbf{y}B = \mathbf{b} \ \text{ in } \ \mathbf{y} \in \Lambda, \tag{5.28}$$

where $\mathbf{b} := \varphi_M(\mathbf{c})$ and

$$B := \begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix}$$

is an integer $Ns \times Ms$-matrix. Notice that $\varphi_M(\mathbf{w}^A) = \varphi_N(\mathbf{w})B$ for $\mathbf{w} \in (\mathcal{O}_S^*)^N$. By assumption, equation (5.28) is solvable, and in view of (5.24), we need to find a solution $\mathbf{y}_0$ of (5.28) such that $\|\mathbf{y}_0\|$ is at most two times the upper bound from Proposition 5.5.

Put $B(\Lambda) := \{\mathbf{y}B : \mathbf{y} \in \Lambda\}$. Clearly, $B(\Lambda)$ is a discrete subgroup in $\mathbb{R}^{Ms}$. Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ be the images of the chosen basis $\mathbf{w}_1, \dots, \mathbf{w}_r$ of Γ modulo Γ_{tors} under φ. Then $\mathbf{v}_1, \dots, \mathbf{v}_r$ form a basis of Λ,

$$\|\mathbf{v}_i\| \le 2h_0 \text{ for } i = 1, \dots, r, \tag{5.29}$$

and $\mathbf{v}_1 B, \dots, \mathbf{v}_r B$ form a system of generators for $B(\Lambda)$. Suppose that the rank of $B(\Lambda)$ is k. By Lemma 5.6 there exists a basis $\mathbf{a}_1, \dots, \mathbf{a}_k$ of $B(\Lambda)$, such that for every $\mathbf{x} = n_1\mathbf{a}_1 + \cdots + n_k\mathbf{a}_k \in B(\Lambda)$ with $n_1, \dots, n_k \in \mathbb{Z}$ we have

$$\|\mathbf{x}\|^* \ge 4^{-k} \max(|n_1|\|\mathbf{a}_1\|^*, \dots, |n_k|\|\mathbf{a}_k\|^*). \tag{5.30}$$

Since $\mathbf{b} \in B(\Lambda)$, there exist $m_1, \dots, m_k \in \mathbb{Z}$, such that $\mathbf{b} = m_1\mathbf{a}_1 + \cdots + m_k\mathbf{a}_k$. Using $\|\mathbf{b}\|^* = 2h(\mathbf{c}) \le 2H$ (in view of (5.24)), $\|\mathbf{a}_i\|^* \ge 2c(d)^{-1}$ (by Lemma 4.1, (5.24) and the fact that $\mathbf{a}_i \in \varphi_M((\mathcal{O}_S^*)^M)$ and (5.30) we have

$$|m_i| \le 4^k \frac{\|\mathbf{b}\|^*}{\|\mathbf{a}_i\|^*} \le 4^k c(d) \cdot H \quad \text{for } i = 1, \dots, k. \tag{5.31}$$

Further, since $\mathbf{v}_i B \in B(\Lambda)$ we can write $\mathbf{v}_i B = \sum_{j=1}^k u_{ij}\mathbf{a}_j$ for $i = 1, \dots, r$. Using $\|\mathbf{v}_i B\|^* = 2h(\mathbf{w}_i^A) \le 2M\delta h_0$ and again $\|\mathbf{a}_j\|^* \ge 2c(d)^{-1}$, (5.30) we get

$$|u_{ij}| \le 4^k c(d) M\delta h_0 \quad \text{for } i = 1, \dots, r, \ j = 1, \dots, k. \tag{5.32}$$

Let $\mathbf{y}$ be a solution of (5.28). Then $\mathbf{y} \in \Lambda$ and so we have $\mathbf{y} = \sum_{i=1}^r \mu_i \mathbf{v}_i$ with $\mu_i \in \mathbb{Z}$ for $i = 1, \dots, r$. Using that on the one hand $\mathbf{b} = m_1\mathbf{a}_1 + \cdots + m_k\mathbf{a}_k$ and on the other hand

$$\mathbf{b} = \mathbf{y}B = \sum_{i=1}^r \mu_i(\mathbf{v}_i B) = \sum_{i=1}^r \mu_i \left(\sum_{j=1}^k u_{ij}\mathbf{a}_j \right) = \sum_{j=1}^k \left(\sum_{i=1}^r u_{ij}\mu_i \right) \mathbf{a}_j,$$

we obtain

$$\sum_{i=1}^r u_{ij}\mu_i = m_j \quad \text{for } j = 1, \dots, k. \tag{5.33}$$

Further we have (5.32) and (5.31) to bound the coefficients and the right hand side of the system of linear equations (5.33). On applying Lemma 5.7 with $V = 4^k c(d) M\delta h_0$, $R = 4^k c(d) H$, we see that the system (5.33) has a solution $\mu \in \mathbb{Z}^r$ with

$$\sum_{i=1}^r |\mu_i| \le r k^{k/2} V^{k-1} \max(V, R) \le \left(r \cdot 4^r c(d) M\delta h_0 \right)^r \cdot H.$$

Now in view of (5.29), the vector $\mathbf{y}_0 = \sum_{i=1}^r \mu_i \mathbf{v}_i$ is a solution to (5.28) such that

$$\|\mathbf{y}_0\| \le 2h_0 \cdot \left(r 4^r c(d) M\delta h_0 \right)^r \cdot H$$

and this is indeed twice the bound of our Proposition. □

As before, Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$ of rank r, A an integer $N \times M$-matrix and $\mathbf{c}$ a point in $(\overline{\mathbb{Q}}^*)^M$. The set $\overline{\Gamma}_\varepsilon$ ($\varepsilon > 0$) is defined as in the Introduction. We assume that A has rank $N - P$.

Proposition 5.8. *Let $\varepsilon > 0$. There exist effectively computable constants C_6, C_7 depending only on $\Gamma, A, \mathbf{c}, \varepsilon$, such that if*

$$\mathbf{x}^A = \mathbf{c} \text{ in } \mathbf{x} \in \overline{\Gamma}_\varepsilon \tag{5.34}$$

is solvable, then there exists $\mathbf{x}_0 \in \overline{\Gamma}_\varepsilon$ with

$$\mathbf{x}_0^A = \mathbf{c}, \quad h(\mathbf{x}_0) \le C_6, \quad [\mathbb{Q}(\mathbf{x}_0) : \mathbb{Q}] \le C_7. \tag{5.35}$$

We deduce Proposition 5.8 from Proposition 5.9 below.

Proposition 5.9. *Let $\mathbf{c}_0 \in (\overline{\mathbb{Q}}^*)^N$, B an integer $P \times N$ matrix of rank P and $\varepsilon > 0$. There exist effectively computable constants C_8, C_9 depending only on $\Gamma, B, \mathbf{c}_0, \varepsilon$, such that if there is $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ with*

$$\mathbf{c}_0 \mathbf{t}^B \in \overline{\Gamma}_\varepsilon, \tag{5.36}$$

then there exists $\mathbf{t}_0 \in (\overline{\mathbb{Q}}^*)^P$ such that

$$\mathbf{c}_0 \mathbf{t}_0^B \in \overline{\Gamma}_\varepsilon, \quad h(\mathbf{t}_0) \le C_8, \quad [\mathbb{Q}(\mathbf{t}_0) : \mathbb{Q}] \le C_9. \tag{5.37}$$

Proposition 5.9 $\Longrightarrow$ Proposition 5.8 Let $A, \mathbf{c}, \varepsilon$ be as in Proposition 5.8. Let $\mathbf{x} \in \overline{\Gamma}_\varepsilon$ with $\mathbf{x}^A = \mathbf{c}$. There are matrices $U_1 \in \mathrm{GL}_N(\mathbb{Z})$, $U_2 \in \mathrm{GL}_M(\mathbb{Z})$ such that

$$U_1 A U_2 = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where D is a non-singular integer $(N-P) \times (N-P)$-matrix. Let $\mathbf{x}^* := \mathbf{x}^{U_1^{-1}}$. Write $\mathbf{x}^* = (\mathbf{s}, \mathbf{t})$ where $\mathbf{s} \in (\overline{\mathbb{Q}}^*)^{N-P}$, $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$. We can decompose $\mathbf{x}^*$ as $(\mathbf{s}, \mathbf{1}) \cdot (\mathbf{1}, \mathbf{t})$, where in the first component $\mathbf{1}$ stands for P ones and in the second component for $N - P$ ones. Notice that $\mathbf{s}^D = \mathbf{c}'$, where $\mathbf{c}'$ consists of the first $N - P$ coordinates of $\mathbf{c}^{U_2}$ and hence $\mathbf{s}^\Delta = \mathbf{c}'^{\Delta D^{-1}}$, where $\Delta = \det D$. This shows that $\mathbf{s}$ belongs to a finite, effectively determinable set depending only on $A, \mathbf{c}$.

Put $\mathbf{c}_0 := (\mathbf{s}, \mathbf{1})^{U_1^{-1}}$, and let B be the matrix consisting of the last P rows of U_1. Then B is a $P \times N$-matrix of rank P. Notice that $\mathbf{c}_0^A = \mathbf{c}$, $BA = 0$ and $\mathbf{c}_0 \mathbf{t}^B = \mathbf{x} \in \overline{\Gamma}_\varepsilon$.

By Proposition 5.9, there is $\mathbf{t}_0 \in (\overline{\mathbb{Q}}^*)^P$ with $\mathbf{c}_0 \mathbf{t}_0^B \in \overline{\Gamma}_\varepsilon$, $h(\mathbf{c}_0) \le C_8$, $[\mathbb{Q}(\mathbf{c}_0) : \mathbb{Q}] \le C_9$, where C_8, C_9 are effectively computable in terms of $B, \mathbf{c}_0, \Gamma, \varepsilon$.

Now put $\mathbf{x}_0 := \mathbf{c}_0\mathbf{t}_0^B$. Then $\mathbf{x}_0 \in \overline{\Gamma}_\varepsilon$, $\mathbf{x}_0^A = \mathbf{c}_0^A\mathbf{t}_0^{BA} = \mathbf{c}$ and $h(\mathbf{x}_0) \le C_6$, $[\mathbb{Q}(\mathbf{x}_0) : \mathbb{Q}] \le C_7$ with C_6, C_7 effectively computable in terms of $B, \mathbf{c}_0, \Gamma, \varepsilon$. Since $\mathbf{c}_0 = (\mathbf{s}, \mathbf{1})^{U_1^{-1}}$ belongs to a finite set effectively computable in terms of $\mathbf{c}, A$ and since B is effectively computable in terms of A, we may choose C_6, C_7 to be effectively computable in terms of $A, \mathbf{c}, \Gamma, \varepsilon$. This proves Proposition 5.8. □

We proceed to prove Proposition 5.9. Let K' be the number field generated by the coordinates of $\mathbf{c}_0$ and by the coordinates of a system of generators for Γ.

Lemma 5.10. *Assume there exists* $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ *with (5.36). Then there exists* $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ *such that*

$$\mathbf{c}_0\mathbf{t}^B \in \overline{\Gamma}_\varepsilon, \quad \exists m \in \mathbb{Z}_{>0} \quad \text{with} \quad \mathbf{t}^m \in (K'^*)^P. \tag{5.38}$$

Proof. First observe that if $\mathbf{u} \in \overline{\Gamma}_\varepsilon$ and $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K')$, then $\sigma(\mathbf{u}) \in \overline{\Gamma}_\varepsilon$. Indeed, write $\mathbf{u} = \mathbf{u}_1\mathbf{u}_2$ with $\mathbf{u}_1 \in \overline{\Gamma}$, $h(\mathbf{u}_2) < \varepsilon$. There is $k \in \mathbb{Z}_{>0}$ such that $\sigma(\mathbf{u}_1)^k = \mathbf{u}_1^k \in \Gamma$, implying that $\sigma(\mathbf{u}_1) \in \overline{\Gamma}$. Further, $h(\sigma(\mathbf{u}_2)) < \varepsilon$. So $\sigma(\mathbf{u}) \in \overline{\Gamma}_\varepsilon$.

Now let $\sigma_1, \dots, \sigma_m$ be the distinct K'-isomorphic embeddings of $K'(\mathbf{t})$ into $\overline{\mathbb{Q}}$. Take

$$\mathbf{t}' := \left(\prod_{i=1}^m \sigma_i(\mathbf{t})\right)^{1/m}.$$

This is determined only up to a factor in $(\overline{\mathbb{Q}}^*_{\mathrm{tors}})^P$, but this is not causing any problem. Write $\mathbf{c}_0\mathbf{t}^B = \mathbf{u}_1\mathbf{u}_2$ with $\mathbf{u}_1 \in \overline{\Gamma}$, $h(\mathbf{u}_2) < \varepsilon$. Then

$$\mathbf{c}_0\mathbf{t}'^B = \left(\prod_{i=1}^m \sigma_i(\mathbf{u}_1)\right)^{1/m} \left(\prod_{i=1}^m \sigma_i(\mathbf{u}_2)\right)^{1/m},$$

which belongs to $\overline{\Gamma}_\varepsilon$. Clearly, $\mathbf{t}^m \in (K'^*)^P$. □

Let S' be the smallest set of places of K', containing all infinite places and such that $\mathbf{c}_0 \in (\mathcal{O}^*_{S'})^N$, $\Gamma \subseteq (\mathcal{O}^*_{S'})^N$. Put $s' := \#S'$.

Lemma 5.11. *Assume there exists* $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ *with (5.36). Then there exists* $\mathbf{t}$ *with*

$$\mathbf{c}_0\mathbf{t}^B \in \overline{\Gamma}_\varepsilon, \quad \mathbf{t} \in (\overline{\mathcal{O}^*_{S'}})^P, \tag{5.39}$$

where $\overline{\mathcal{O}^*_{S'}} = \{\mathbf{x} \in \overline{\mathbb{Q}}^* : \exists m \in \mathbb{Z}_{>0} \text{ with } \mathbf{x}^m \in \mathcal{O}^*_{S'}\}$.

Proof. Let $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ be as in (5.38), i.e., $\mathbf{t}^m \in (K'^*)^P$ for some $m \in \mathbb{Z}_{>0}$. Write

$$\mathbf{c}_0 \mathbf{t}^B = \mathbf{y}\mathbf{z} \quad \text{with} \quad \mathbf{y} \in \overline{\Gamma}, \quad h(\mathbf{z}) < \varepsilon. \tag{5.40}$$

Let $n \in \mathbb{Z}_{>0}$ be such that $\mathbf{y}^n \in \Gamma$ and let k be any positive multiple of $\operatorname{lcm}(m, n)$. Thus

$$\mathbf{z}^k = \mathbf{c}_0^k (\mathbf{t}^k)^B \mathbf{y}^{-k} \in (K'^*)^P. \tag{5.41}$$

Write $\mathbf{t} = (t_1, \dots, t_P)$. By the Dirichlet-Chevalley-Weil S-unit theorem, there are $\varepsilon_1, \dots, \varepsilon_P \in \mathcal{O}_{S'}^*$ such that

$$\left| \log |\varepsilon_i|_v - \log \left(\frac{|t_i^k|_v}{\left\{ \prod_{w \in S} |t_i^k|_w \right\}^{1/s}} \right) \right| \le C \quad \text{for} \quad i = 1, \dots, P, \ v \in S',$$

where C is an effectively computable constant depending only on K', S', and independent of k. Now define

$$\begin{aligned} &\mathbf{t}' := (\varepsilon_1^{1/k}, \dots, \varepsilon_P^{1/k}), \quad \mathbf{z}' := \mathbf{c}_0 (\mathbf{t}'^B) \mathbf{y}^{-1}, \\ &\boldsymbol{\eta} = (\eta_1, \dots, \eta_N) := (\varepsilon_1, \dots, \varepsilon_P)^B. \end{aligned} \tag{5.42}$$

(with a suitable choice of the k-th roots). Write $\mathbf{z} := (z_1, \dots, z_P)$, $\mathbf{z}' := (z_1', \dots, z_P')$, $\mathbf{v} := (v_1, \dots, v_N) = \mathbf{t}^B$, $(\mathbf{c}_0 \mathbf{y}^{-1})^k := (\alpha_1, \dots, \alpha_P)$. Then since $\alpha_1, \dots, \alpha_P \in \mathcal{O}_{S'}^*$ (by our choice of k and S') we have for $i = 1, \dots, N$, $v \in S'$,

$$\begin{aligned} &\left| \log |z_i'^k|_v - \log \left(\frac{|z_i^k|_v}{\left\{ \prod_{w \in S'} |z_i^k|_w \right\}^{1/s'}} \right) \right| \\ &\quad = \left| \log |\alpha_i \eta_i|_v - \log \left(\frac{|\alpha_i v_i^k|_v}{\left\{ \prod_{w \in S'} |\alpha_i v_i^k|_w \right\}^{1/s'}} \right) \right| \\ &\quad = \left| \log |\eta_i|_v - \log \left(\frac{|v_i^k|_v}{\left\{ \prod_{w \in S'} |v_i^k|_w \right\}^{1/s'}} \right) \right| \le C', \end{aligned}$$

where C' is an effectively computable constant depending only on K', S' and B, but which is independent of k. Together with the product formula this implies

$$\left| \log |z_i'^k|_v - \log |z_i^k|_v - \frac{1}{s'} \sum_{w \notin S'} \log |z_i^k|_w \right| \le C'$$

for $v \in S'$, $i = 1, \dots, N$. Now we get

$$\begin{aligned} h(\mathbf{z}'^k) &= \sum_{i=1}^{N} \sum_{v \in S'} \max\left(0, \log |z_i'^k|_v\right) \\ &\le \sum_{i=1}^{N} \sum_{v \in S'} \max\left(0, C' + \log |{z_i}^k|_v + \frac{1}{s'} \sum_{w \notin S'} \log |z_i^k|_w\right) \\ &\le Ns'C' + \sum_{i=1}^{N} \sum_{v \in S'} \max\left(0, \log |{z_i}^k|_v\right) + \sum_{i=1}^{N} \sum_{v \notin S'} \max\left(0, \log |{z_i}^k|_v\right) \\ &= Ns'C' + \sum_{i=1}^{N} h(z_i^k) = Ns'C' + h(\mathbf{z}^k). \end{aligned}$$

Consequently,

$$h(\mathbf{z}') \le h(\mathbf{z}) + \frac{Ns'C'}{k}.$$

By assumption, $h(\mathbf{z}) < \varepsilon$. We had chosen k to be any positive multiple of $\operatorname{lcm}(m, n)$. By choosing k large enough, we can achieve that $h(\mathbf{z}') < \varepsilon$. Now from our choice of $\mathbf{t}'$ in (5.42) it follows that $\mathbf{t}' \in (\overline{\mathcal{O}_S^*})^P$ and $\mathbf{c}_0 \mathbf{t}'^B = \mathbf{y}\mathbf{z}' \in \overline{\Gamma}_\varepsilon$. This proves Lemma 5.11. □

The proof of Proposition 5.9 rests upon linear programming.

Define the group

$$G := \left\{ \mathbf{y}\mathbf{t}^B \ : \ \mathbf{y} \in \overline{\Gamma}, \mathbf{t} \in (\overline{\mathcal{O}_{S'}^*})^P \right\}.$$

This is a group of finite rank q. Choose a maximal multiplicatively independent subset $\mathbf{t}_1, \dots, \mathbf{t}_s$ of $(\mathcal{O}_{S'}^*)^P$. Then $\mathbf{u}_1 := \mathbf{t}_1^B, \dots, \mathbf{u}_s := \mathbf{t}_s^B$ are multiplicatively independent since $\operatorname{rank} B = P$. Choose $\mathbf{u}_{s+1}, \dots, \mathbf{u}_q \in \Gamma$ such that $\{\mathbf{u}_1, \dots, \mathbf{u}_q\}$ form a maximal multiplicatively independent subset of G. After a suitable choice of roots of $\mathbf{u}_1, \dots, \mathbf{u}_q$, we may express G as

$$G = \left\{ \boldsymbol{\rho} \mathbf{u}_1^{\xi_1} \dots \mathbf{u}_q^{\xi_q} \ : \ \boldsymbol{\rho} \in (\overline{\mathbb{Q}}_{\text{tors}}^*)^N, \xi_1, \dots, \xi_q \in \mathbb{Q} \right\}.$$

We are searching for $\mathbf{t} \in (\overline{\mathcal{O}_{S'}}^*)^P$ such that

$$\mathbf{c}_0 \mathbf{t}^B = \mathbf{y}\mathbf{z}, \quad \mathbf{y} \in \overline{\Gamma}, \quad h(\mathbf{z}) < \varepsilon.$$

For such $\mathbf{t}$ we have $\mathbf{z} = \mathbf{c}_0 \mathbf{y}^{-1} \mathbf{t}^B \in \mathbf{c}_0 G$. So we are searching for $\mathbf{z} \in \mathbf{c}_0 G$ with $h(\mathbf{z}) < \varepsilon$.

We give an expression for the height of an element $\mathbf{z} \in \mathbf{c}_0 G$. Such an element can be expressed as

$$\mathbf{z} = \mathbf{c}_0 \boldsymbol{\rho} \mathbf{u}_1^{\xi_1} \dots \mathbf{u}_q^{\xi_q} \ \text{ with } \ \boldsymbol{\rho} \in (\overline{\mathbb{Q}}_{\text{tors}}^*)^N, \xi_1, \dots, \xi_q \in \mathbb{Q}.$$

Write $\boldsymbol{\xi} := (\xi_1, \dots, \xi_q)$. Let k be a positive integer such that $\boldsymbol{\rho}^k = \mathbf{1}$, $k\boldsymbol{\xi} \in \mathbb{Z}^q$. Further, write $\mathbf{u}_i = (u_{i1}, \dots, u_{iN})$ $(i = 1, \dots, r)$, $\mathbf{c}_0 = (c_{01}, \dots, c_{0N})$. Then

$$\begin{aligned} h(\mathbf{z}) = \frac{1}{k} h(\mathbf{z}^k) &= \frac{1}{k} \sum_{i=1}^{N} \sum_{v \in S'} \max \left(0, k \log |c_{0i}|_v + \sum_{j=1}^{q} k\xi_j \log |u_{ij}|_v \right) \\ &= \sum_{i=1}^{N} \sum_{v \in S'} \max \left(0, \log |c_{0i}|_v + \sum_{j=1}^{q} \xi_j \log |u_{ij}|_v \right) \qquad (5.43) \\ &= \frac{1}{2} \sum_{i=1}^{N} \sum_{v \in S'} \left| \log |c_{0i}|_v + \sum_{j=1}^{q} \xi_j \log |u_{ij}|_v \right| =: f(\boldsymbol{\xi}), \end{aligned}$$

where we have used $\sum_{v \in S'} \log |c_{0i}|_v = 0$, $\sum_{v \in S'} \log |u_{ij}|_v = 0$ for all i, j. The function f can be extended to $\mathbb{R}^q$. We prove some properties of this function.

Lemma 5.12. (i) *For every $R \geq 0$, the set $\{\boldsymbol{\xi} \in \mathbb{R}^q : f(\boldsymbol{\xi}) \leq R\}$ is compact with respect to the topology in $\mathbb{R}^q$.*
(ii) *There is an effectively computable constant $C > 0$ such that*
$|f(\boldsymbol{\xi}_1) - f(\boldsymbol{\xi}_2)| \leq C \|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|_\infty$ for all $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathbb{R}^q$.

Proof. (i). We can express $f(\boldsymbol{\xi})$ as $\|\alpha(\boldsymbol{\xi})\|$ where $\|\cdot\|$ is a norm on $\mathbb{R}^{Ns}$ and α an injective affine map from $\mathbb{R}^q$ to $\mathbb{R}^{Ns}$. So our set under consideration is homeomorphic to a closed subset of a compact set, hence compact.

(ii). Obvious. □

Lemma 5.13. *The function f assumes a minimum on $\mathbb{R}^r$ and it is possible to determine effectively*

$$\varepsilon_0 := \min\{f(\boldsymbol{\xi}) \ : \ \boldsymbol{\xi} \in \mathbb{R}^q\}$$

and $\boldsymbol{\xi}_0$ with $f(\boldsymbol{\xi}_0) = \varepsilon_0$.

Proof. It clearly suffices to prove that f assumes a minimum on

$$D := \{\boldsymbol{\xi} \in \mathbb{R}^q \ : \ f(\boldsymbol{\xi}) \leq f(\mathbf{0})\}$$

and to determine the minimum of f on D and a point in D where this minimum is assumed. By Lemma 5.12,(i) the set D is compact, so f does indeed assume its minimum on D.

We can rewrite f as

$$f(\boldsymbol{\xi}) = \max\left(L_1(\boldsymbol{\xi}) + \beta_1, \dots, L_A(\boldsymbol{\xi}) + \beta_A\right),$$

where $L_1, \dots, L_A$ are linear forms with real coefficients and $\beta_1, \dots, \beta_A \in \mathbb{R}$. For $i = 1, \dots, A$, let

$$D_i := \{\boldsymbol{\xi} \in D : \ L_i(\boldsymbol{\xi}) + \beta_i \geq L_j(\boldsymbol{\xi}) + \beta_j \ \text{ for } \ j = 1, \dots, A, j \neq i\}.$$

The set D_i is a closed subset of D, hence compact. Thus D_i is a compact polytope. Notice that $f(\boldsymbol{\xi}) = L_i(\boldsymbol{\xi}) + \beta_i$ for $\boldsymbol{\xi} \in D_i$. ¿From the theory of linear programming it follows that f assumes its minimum on D_i in a vertex of D_i. The vertices of D_i can be determined effectively. So we can effectively determine $\varepsilon_i := \min\{L_i(\boldsymbol{\xi}) + \beta_i \ : \ \boldsymbol{\xi} \in \mathcal{D}_i\}$ and $\boldsymbol{\xi}_i \in D_i$ with $f(\boldsymbol{\xi}_\mathbf{i}) = \varepsilon_i$.

Now $\varepsilon_0 = \min(\varepsilon_1, \dots, \varepsilon_A)$, and $f(\boldsymbol{\xi}_0) = \varepsilon_0$, where $\boldsymbol{\xi}_0$ is the point $\boldsymbol{\xi}_i$ among $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_A$ such that $\varepsilon_i = \varepsilon_0$. □

Proof of Proposition 5.9. Assume that there exists $\mathbf{t} \in (\overline{\mathcal{O}^*_{S'}})^P$ such that $\mathbf{c}_0\mathbf{t}^B \in \overline{\Gamma}_\varepsilon$. Write $\mathbf{c}_0\mathbf{t}^B = \mathbf{yz}$ with $\mathbf{y} \in \overline{\Gamma}$, $h(\mathbf{z}) < \varepsilon$. Then $\mathbf{z} = \mathbf{c}_0\mathbf{y}^{-1}\mathbf{t}^B = \mathbf{c}_0\boldsymbol{\rho}\mathbf{u}_1^{\xi_1} \dots \mathbf{u}_q^{\xi_q}$ with $\boldsymbol{\rho} \in (\overline{\mathbb{Q}}^*_{\text{tors}})^N$ and $\xi_1, \dots, \xi_q \in \mathbb{Q}$. So $h(\mathbf{z}) = f(\boldsymbol{\xi}) < \varepsilon$ and therefore, $\varepsilon > \varepsilon_0$. Let C be the constant from Lemma 5.12,(ii) and define the integer k by

$$k := \left[\frac{2C}{\varepsilon - \varepsilon_0}\right] + 1. \tag{5.44}$$

Let $\boldsymbol{\xi}_0$ be as in Lemma 5.13 and write $\boldsymbol{\xi}_0 = (\xi_{01}, \dots, \xi_{0q})$. Define integers $n_1, \dots, n_q$ by

$$|k\xi_{0i} - n_i| < 1 \ \ (i = 1, \dots, q)$$

and let

$$\mathbf{t}_0 = \mathbf{t}_1^{n_1/k} \dots \mathbf{t}_s^{n_s/k}, \ \ \mathbf{z}_0 = \mathbf{c}_0\mathbf{u}_1^{n_1/k} \dots \mathbf{u}_q^{n_q/k}.$$

By Lemma 5.12,(ii) and (5.44),

$$\begin{aligned} h(\mathbf{z}_0) = f\left(\frac{n_1}{k}, \dots, \frac{n_q}{k}\right) &\leq f(\boldsymbol{\xi}_0) + C \max_{1 \leq i \leq q} \left|\xi_{0i} - \frac{n_i}{k}\right| \\ &\leq \varepsilon_0 + \frac{C}{k} < \varepsilon_0 + \frac{C(\varepsilon - \varepsilon_0)}{2C} < \varepsilon. \end{aligned}$$

Further,

$$\begin{aligned} &\mathbf{c}_0\mathbf{t}_0^B = \mathbf{u}_{s+1}^{-n_{s+1}/k} \dots \mathbf{u}_q^{-n_q/k}\mathbf{z}_0 \in \overline{\Gamma}_\varepsilon, \\ &h(\mathbf{t}_0) \leq \sum_{i=1}^{s} \left|\frac{n_i}{k}\right| h(\mathbf{t}_i) \leq \sum_{i=1}^{q} (|\xi_{0i}| + \frac{1}{k}) h(\mathbf{t}_i) \leq C_8 \end{aligned}$$

and $\mathbf{t}_0^k \in (K'^*)^P$, implying $[\mathbb{Q}(\mathbf{t}_0) : \mathbb{Q}] \leq C_9$. The quantities C, ε_0, as well as $\mathbf{t}_1, \dots, \mathbf{t}_s$ are effectively computable in terms of $\Gamma, B, \mathbf{c}_0$, while k is effectively computable in terms of these parameters and also ε. Hence the constants C_8, C_9 are indeed effectively computable in terms of $\Gamma, B, \mathbf{c}_0, \varepsilon$, but they have been defined only for $\varepsilon > \varepsilon_0$. For completeness, we define $C_8 := 1$, $C_9 := 1$ if $\varepsilon \leq \varepsilon_0$. Then clearly, Proposition 5.9 holds with these C_8, C_9. □

5.6 Proof of Theorem 2.11.

We write $\mathbf{x}^{\mathbf{a}} := x_1^{a_1} \cdots x_N^{a_N}$ for $\mathbf{x} = (x_1, \ldots, x_N) \in (\overline{\mathbb{Q}}^*)^N$, $\mathbf{a} = (a_1, \ldots, a_N) \in \mathbb{Z}^N$.

By assumption

$$\mathcal{X} = \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \ : \ f_1(\mathbf{x}) = 0, \ldots, f_m(\mathbf{x}) = 0 \right\},$$

where each polynomial f_i belongs to $\overline{\mathbb{Q}}[X_1, \ldots, X_N]$ and has at most three non-zero terms. Further, $\deg f_i \le \delta$ and $\max(1, h(f_i)) \le H$ for $i = 1, \ldots, m$. Without loss of generality we assume that f_i $(i = 1, \ldots, n)$ are trinomials and f_i $(i = n+1, \ldots, m)$ are binomials, where $0 \le n \le m$. Thus, by dividing each f_i by one of its terms we see that $\mathcal{X}$ is given by equations

$$\alpha_{i1}\mathbf{x}^{\mathbf{a}_{i1}} + \alpha_{i2}\mathbf{x}^{\mathbf{a}_{i2}} = 1 \ \ (i = 1, \ldots, n), \qquad \alpha_{i1}\mathbf{x}^{\mathbf{a}_{i1}} = 1 \ \ (i = n+1, \ldots, m), \tag{5.45}$$

where $\alpha_{ij} \in \overline{\mathbb{Q}}^*$, $\mathbf{a}_{ij} \in \mathbb{Z}^N$ for $(i, j) \in I := \{(1,1), \ldots, (m,1), (1,2), \ldots, (n,2)\}$. We observe here that since each polynomial f_i has total degree at most δ, we have estimates for the maximum norm and the sum norm,

$$\|\mathbf{a}_{ij}\|_\infty \le \delta, \quad \|\mathbf{a}_{ij}\|_1 \le 2\delta \ \text{ for } (i,j) \in I. \tag{5.46}$$

Clearly the stabilizer of $\mathcal{X}$ is given by

$$\begin{aligned} \mathcal{H} := \mathrm{Stab}(\mathcal{X}) &= \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \ \mid \mathbf{x}^{\mathbf{a}_{ij}} = 1 \text{ for } i = 1, \ldots, m, \ j = 1, 2 \right\} \\ &= \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \ \mid \mathbf{x}^A = \mathbf{1} \right\}, \end{aligned} \tag{5.47}$$

where A is the $N \times (2m - n)$ matrix with columns $\mathbf{a}_{ij}$ $((i,j) \in I)$.

Let $i \in \{1, \ldots, n\}$. Let $\mathbf{x} \in \mathcal{X} \cap \Gamma$. Denote by G be the subgroup of K^* generated by $\xi_1 := \mathbf{w}_1^{a_{i1}}, \ldots, \xi_r := \mathbf{w}_r^{a_{i1}}$. Then for the quantity Q defined by (5.4), we have

$$Q := \prod_{j=1}^{r} \max\left(1, h(\mathbf{w}_1^{a_{ij}})\right) \le (\delta h_0)^r.$$

We apply Lemma 5.2 to the equation $\alpha_{i1}x + \alpha_{i2}y = 1$ in $x \in G$, $y \in \mathcal{O}_S^*$. Notice that $\max(1, h(\alpha_{i1}), h(\alpha_{i2})) \le H$. Replacing Q by $(\delta h_0)^r$ in the expression for C_5, we obtain a constant bounded above by C^*. In fact, this can be shown by a straightforward computation, using that the term with the maximum in C_5 is bounded above by $46r^2 \log^* \max(ds\mathbf{P}, \delta h_0)$. It follows that

$$h(\mathbf{x}^{\mathbf{a}_{i,j}}) < C^* H \text{ for } i = 1, \ldots, n, \ j = 1, 2. \tag{5.48}$$

We clearly also have $h(\mathbf{x}^{\mathbf{a}_{i1}}) = h(\alpha_{i1}^{-1}) \le H$ for $i = n+1, \dots, m$. So we have (5.48) for $(i,j) \in I$. This implies

$$\mathbf{x}^A = \mathbf{c}, \tag{5.49}$$

where A is the $N \times (2m-n)$-matrix from above and where $\mathbf{c} \in (K^*)^{2m-n}$ with

$$h(\mathbf{c}) \le (2m-n)C^*H \le 2mC^*H. \tag{5.50}$$

Further, the entries of A have absolute values at most δ, and of each column of A the sum of its absolute values is at most 2δ.

We first assume that the stabilizer $\mathcal{H}$ is finite. Then A has rank N. Suppose for convenience that the first N columns, $\mathbf{a}_1, \dots, \mathbf{a}_N$, say, of A form an invertible matrix D, with determinant Δ. Let $\mathbf{c}'$ consist of the first N coordinates of $\mathbf{c}$. Then $\mathbf{x}^{\Delta} = \mathbf{c}'^{\Delta D^{-1}}$. By Hadamard's inequality and (5.46), the entries of ΔD^{-1} have absolute value at most

$$\max_{1 \le i \le N} \prod_{j \ne i} \|\mathbf{a}_j\|_2 \le (2\delta)^{N-1}. \tag{5.51}$$

Further, $h(\mathbf{c}') \le NC^*H$. So $h(\mathbf{x}) \le N(2\delta)^{N-1}C^*H = C_2H$. This proves part (i).

We now assume that $\mathcal{H}$ is infinite. Notice that we have to consider finitely many systems (5.49) as $\mathbf{c}$ runs through a finite set. If such a system has a solution $\mathbf{x}$ with $\mathbf{x} \in \mathcal{X}$, then each element of the translate $\mathbf{x}\mathcal{H}$ is also a solution of this system. On the other hand $\mathbf{x}\mathcal{H} \subset \mathcal{X}$. Thus we have proved that $\mathcal{X} \cap \Gamma$ is contained in some finite union of translates

$$\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H}$$

with $\mathbf{x}_i\mathcal{H} \subset \mathcal{X}$ for $i = 1, \dots, T$.

Fix any of these translates, which means that we have fixed one of the systems from (5.49). By assumption this system has a solution in $\mathbf{x} \in \Gamma$. Now by Proposition 5.5 (with $M = 2m-n \le 2m$) and (5.50), this fixed system of type (5.49) has a solution $\mathbf{x} \in \Gamma$ such that

$$h(\mathbf{x}) \le h_0\big(2r4^r c(d) m\delta h_0\big)^r \cdot 2mC^*H \le C_3H.$$

This proves Theorem 2.11.

5.7 Proofs of Theorems 2.12 and 2.13

Proof of Theorem 2.13. Let $\mathbf{x} \in \mathcal{X}(\overline{\mathbb{Q}}) \cap C(\overline{\Gamma}, \varepsilon)$, with the value of ε given in (2.35).

As before, we write $\mathbf{x} = \mathbf{yz}$ with $\mathbf{y} \in \overline{\Gamma}$ and $\mathbf{z} \in (\overline{\mathbb{Q}}^*)^2$ with $h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y}))$ and we may further split up $\mathbf{y}$ as $\mathbf{y} = \mathbf{vw}$ with $\mathbf{v} \in \Gamma$, $\mathbf{w} = \prod_{i=1}^{r} \mathbf{w}_i^{\gamma_i}$, where $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \le \frac{1}{2}$.

Define new polynomials $f_i^*(\mathbf{V}) := f_i(\mathbf{wz}\cdot\mathbf{V})$ $(i=1,\dots,m)$ and let $\mathcal{X}^*$ be the variety given by $f_i^*=0$ for $i=1,\dots,m$, i.e., $\mathcal{X}^* := (\mathbf{wz})^{-1}\mathcal{X}$. Then $\mathbf{v}\in\mathcal{X}^*\cap\Gamma$. Notice that $\deg f_i^*\le\delta$, and $\max(1,h(f_i^*))\le H+\delta h(\mathbf{wz})\le H+\delta(h(\mathbf{w})+h(\mathbf{z}))$ for $i=1,\dots,m$.

We observe that $\mathcal{X}$ and $\mathcal{X}^*$ have the same stabilizer $\mathcal{H}$, and this stabilizer is assumed to be finite.

We obtain the upper bound for $h(\mathbf{x})$ by applying Theorem 2.11 to $\mathcal{X}^*$ and then following the proof of Theorem 2.10, replacing everywhere C_1 by C_2.

Now we estimate $[L(\mathbf{x}):L]$. To this end, it suffices to estimate the number of distinct points among $\sigma(\mathbf{x})$, $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/L)$.

Let $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Write again $\mathbf{x}=\mathbf{yz}$ such that $\mathbf{y}\in\overline{\Gamma}$, $h(\mathbf{z})<\varepsilon(1+h(\mathbf{y}))$. Put $\mathbf{u}_\sigma := \sigma(\mathbf{x})\mathbf{x}^{-1}$. Since $\sigma(\mathbf{y})\mathbf{y}^{-1}$ is a torsion point, we have

$$h(\mathbf{u}_\sigma)=h(\sigma(\mathbf{x})\mathbf{x}^{-1})=h(\sigma(\mathbf{z})\mathbf{z}^{-1})\le 2h(\mathbf{z}).$$

Completely similarly as (5.23) we obtain

$$h(\mathbf{z})\le\varepsilon\cdot\big(C_2\delta r h_0+2C_2H\big). \tag{5.52}$$

Hence

$$h(\mathbf{u}_\sigma)\le 2\varepsilon\cdot\big(C_2\delta r h_0+2C_2H\big)=:\eta.$$

We assume again that f_i is a trinomial for $i=1,\dots,n$ and a binomial for $i=n+1,\dots,m$. Then (5.45) holds for certain integer vectors $\mathbf{a}_{ij}$ and we obtain

$$\tilde\alpha_{i1}\mathbf{u}_\sigma^{\mathbf{a}_{i1}}+\tilde\alpha_{i2}\mathbf{u}_\sigma^{\mathbf{a}_{i2}}=1 \text{ for } i=1,\dots,n,$$
$$\tilde\alpha_{i1}\mathbf{u}_\sigma^{\mathbf{a}_{i1}}=1 \text{ for } i=n+1,\dots,m.$$

Let $i\in\{1,\dots,n\}$. By our choice of ε in (2.35) we have

$$h(\mathbf{u}_\sigma^{\mathbf{a}_{i1}},\mathbf{u}_\sigma^{\mathbf{a}_{i2}})\le 2\delta\eta=0.03.$$

Thus by Lemma 5.3 (i) we see that there are at most 2 possibilities for each pair $(\mathbf{u}_\sigma^{\mathbf{a}_{i1}},\mathbf{u}_\sigma^{\mathbf{a}_{i2}})$. These facts imply that $\mathbf{u}_\sigma^A=\mathbf{c}_\sigma$ where $\mathbf{c}_\sigma$ runs through a set of cardinality at most 2^m if σ runs through $\mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Fix $\mathbf{c}_0$ and then σ_0 with $\mathbf{u}_{\sigma_0}^A=\mathbf{c}_0$. Then for every $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/L)$ with $\mathbf{u}_\sigma^A=\mathbf{c}_0$ we have $\left(\frac{\mathbf{u}_\sigma}{\mathbf{u}_{\sigma_0}}\right)^A=\mathbf{1}$. This shows that for fixed $\mathbf{c}_0$ we have at most $t:=\#\mathcal{H}$ possibilities for $\mathbf{u}_\sigma$, where $\mathcal{H}:=\mathrm{Stab}(\mathcal{X})=\Big\{\mathbf{x}\in(\overline{\mathbb{Q}}^*)^N\mid\mathbf{x}^A=\mathbf{1}\Big\}$. Hence for $\mathbf{u}_\sigma$ we have altogether at most $2^m t$ possibilities, implying $[L(\mathbf{x}):L]\le 2^m t$.

It remains to estimate $t=\#\mathcal{H}$. By assumption, $\mathcal{H}$ is finite hence is zero-dimensional, therefore the matrix A has rank N. Suppose for instance that the first N columns of A

form an invertible matrix D. Then $\mathcal{H}$ is contained in $\mathcal{H}' = \{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \mathbf{x}^D = \mathbf{1}\}$. There are matrices $U_1 \in \mathrm{GL}_N(\mathbb{Z})$, $U_2 \in \mathrm{GL}_M(\mathbb{Z})$ such that $U_1 D U_2 = D_0$ is a diagonal matrix with positive integers $d_1, \dots, d_N$ on the diagonal. Now $\mathbf{x} \mapsto \mathbf{x}^{U_1^{-1}}$ maps $\mathcal{H}'$ isomorphically to the algebraic group given by $x_1^{d_1} = 1, \dots, x_N^{d_N} = 1$ and the latter clearly has cardinality $d_1 \cdots d_N$.

By an estimate similar to (5.51), using (5.46), we have $d_1 \cdots d_N = |\det D| \le (2\delta)^N$. Hence $t \le (2\delta)^N$. This leads to $[L(\mathbf{x}) : L] \le 2^m t \le 2^{m+N}\delta^N$. □

Proof of Theorem 2.12. First suppose that $\mathrm{Stab}(\mathcal{X})$ is finite. Let $\mathbf{x} \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$. We write $\mathbf{x} = \mathbf{yz}$ with $\mathbf{y} \in \overline{\Gamma}$, $h(\mathbf{z}) < \varepsilon$ and then as usual $\mathbf{y} = \mathbf{vw}$ with $\mathbf{v} \in \Gamma$ and $\mathbf{w} = \prod_{i=1}^r \mathbf{w}_i^{\gamma_i}$, where $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \le \frac{1}{2}$. Like in the proof of Theorem 2.13, we define the polynomials $f_i^*(\mathbf{V}) = f_i(\mathbf{wz}\cdot\mathbf{V})$ $(i = 1, \dots, m)$ and let $\mathcal{X}^*$ be the variety given by $f_i^* = 0$ $(i = 1, \dots, m)$. Then again, $\mathbf{v} \in \mathcal{X}^* \cap \Gamma$. Recall that $\deg f_i^* \le \delta$, and that

$$\max(1, h(f_i^*)) \le H + \delta h(\mathbf{wz}) \le H + \delta(h(\mathbf{w}) + h(\mathbf{z})) \le H + \delta\left(\frac{rh_0}{2} + \varepsilon\right)$$

for $i = 1, \dots, m$. Now applying part (i) of Theorem 2.11 to $\mathcal{X}^* = (\mathbf{wz})^{-1}\mathcal{X}$, we obtain

$$h(\mathbf{v}) \le C_2\left(H + \delta\frac{rh_0}{2} + \varepsilon\right)$$

and together with $h(\mathbf{x}) \le h(\mathbf{v}) + h(\mathbf{w}) + h(\mathbf{z}) \le h(\mathbf{v}) + \frac{rh_0}{2} + \varepsilon$ this leads to the upper bound for $h(\mathbf{x})$ in (2.34).

As for the estimation of $[L(\mathbf{x}) : L]$, instead of (5.52) we have $h(\mathbf{z}) < \varepsilon$, then our assumption $\varepsilon = \frac{0.03}{4\delta}$ leads to the same conclusion $h(\mathbf{u}_\sigma^{\mathbf{a}_{i1}}, \mathbf{u}_\sigma^{\mathbf{a}_{i2}}) \le 0.03$ for $i = 1, \dots, n$, and the proof is concluded in the same way as that of Theorem 2.13.

We now assume that $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$ is infinite. We define $\mathbf{z}, \mathbf{v}, \mathbf{w}$ as above and keep the notation from the proof of Theorem 2.11. Thus we obtain, completely similarly as in (5.48),

$$h(\mathbf{v}^{\mathbf{a}_{i,j}}) < C^*\left(H + \delta\frac{rh_0}{2} + \delta\varepsilon\right) \text{ for } (i,j) \in I,$$

and together with $h(\mathbf{w}) \le \frac{rh_0}{2}$, $h(\mathbf{z}) < \varepsilon$, this leads to

$$h(\mathbf{x}^{\mathbf{a}_{i,j}}) < C^*(H + \delta r h_0) \text{ for } (i,j) \in I.$$

Then, similarly as (5.50) we obtain,

$$\mathbf{x}^A = \mathbf{c} \text{ with } h(\mathbf{c}) \le 2mC^*(H + \delta r h_0). \tag{5.53}$$

Let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$, and put $\mathbf{u}_\sigma := \sigma(x) \cdot \mathbf{x}^{-1}$, $\mathbf{c}_\sigma := \sigma(\mathbf{c}) \cdot \mathbf{c}^{-1}$. Thus, $\mathbf{u}_\sigma^A = \mathbf{c}_\sigma$. Following the argument in the proof of Theorem 2.13, using our choice $\varepsilon = \frac{0.03}{4\delta}$ for ε, we infer again that $h(\mathbf{u}_\sigma^{\mathbf{a}_{i1}}, \mathbf{u}_\sigma^{\mathbf{a}_{i2}}) \le 0.03$ for $i = 1, \dots, n$, and subsequently, that $\mathbf{c}_\sigma$ runs through a set of cardinality at most 2^m if σ runs through $\mathrm{Gal}(\overline{\mathbb{Q}}/L)$. This implies that we have at most 2^m possibilities for $\sigma(\mathbf{c})$. Hence

$$[L(\mathbf{c}) : L] \le 2^m. \tag{5.54}$$

Now from (5.53), (5.54) we infer that for every $\mathbf{x} \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$ there is $\mathbf{c}$ from a finite, effectively determinable set depending only on Γ and $f_1, \dots, , f_m$, such that $\mathbf{x}^A = \mathbf{c}$. We conclude by applying Proposition 5.8 to each of the equations $\mathbf{x}^A = \mathbf{c}$. □

Chapter 6

General description of the method for proving effective results over finitely generated domains

In this chapter we collect the main ingredients of the method of Evertse and Győry [32] for proving effective results for Diophantine problems over arbitrary finitely generated domains A containing $\mathbb{Z}$.

6.1 Extending the domain A

In this section we extend the domain A to a larger finitely generated domain B in which it will be more convenient to do effective computations, and which can be chosen in such a way that several elements of K (chosen according to our needs) will be units in this extended domain. This latter property will have special importance when we define our specializations. We also introduce a new representation for elements of K, which gives rise to a different way of measuring elements of K than the one using the size of representatives, and this way of measuring will be more convenient in our proofs.

Let $A = \mathbb{Z}[z_1, \dots, z_r]$ be a finitely generated domain given by (3.1), and let K be its quotient field. Let $q \geq 0$ denote the transcendence degree of K. We may assume without loss of generality that $z_1, \dots, z_q$ is a transcendence basis of $K/\mathbb{Q}$. Put

$$K_0 := \mathbb{Q}(z_1, \dots, z_q), \qquad A_0 := \mathbb{Z}[z_1, \dots, z_q]. \tag{6.1}$$

For elements $f \in A_0 \setminus \{0\}$ let $\deg f$ and $h(f)$ denote the total degree and logarithmic height of f, respectively, viewed as a polynomial in the unknowns $z_1, \dots, z_q$. In the case $q = 0$ we

define $\deg f := 0$ and $h(f) := \log |f|$. Put

$$d_0 := \max(1, \deg f_1, \dots, \deg f_t), \qquad h_0 := \max(1, h(f_1), \dots, h(f_t)), \tag{6.2}$$

where $f_1, \dots, f_t$ are the generators of the ideal $\mathcal{I}$ in (3.1).

Since the field K is a finite algebraic extension of K_0, we can write $K = K_0(w)$ for some $w \in K$. We recall the following result of Evertse and Győry.

Proposition 6.1. *(i) There exists $w \in A$ such that $K = K_0(w)$, w is integral over A_0 and w has minimal polynomial*

$$\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X]$$

over K_0 such that $D \le d_0^{r-q}$ and

$$\deg \mathcal{F}_k \le (2d_0)^{\exp O(r)}, \qquad h(\mathcal{F}_k) \le (2d_0)^{\exp O(r)}(h_0 + 1) \tag{6.3}$$

for $k = 1, \dots, D$.

(ii) Let $\alpha_1, \dots, \alpha_k \in K^*$ and suppose that α_i has representation pair (u_i, v_i) with $u_i, v_i \in \mathbb{Z}[X_1, \dots, X_r]$, $v_i \notin I$, for $i = 1, \dots, k$. Put

$$\begin{aligned} d^{**} &:= \max(d_0, \deg u_1, \deg v_1, \dots, \deg u_k, \deg v_k), \\ h^{**} &:= \max(h_0, h(u_1), h(v_1), \dots, h(u_k), h(v_k)). \end{aligned}$$

Then there is a non-zero $f \in A_0$ such that with $B := A_0[w, f^{-1}]$ we have

$$\begin{aligned} &A \subseteq B, \\ &\alpha_1, \dots, \alpha_k \in B^*. \end{aligned} \tag{6.4}$$

Further, f can be chosen such that it fulfils

$$\deg f \le (k+1)(2d^{**})^{\exp O(r)}, \qquad h(f) \le (k+1)(2d^{**})^{\exp O(r)}(h^{**} + 1). \tag{6.5}$$

Proof. These versions of (i) and (ii) are stated in Proposition 3.1 in [10], however originally (i) is proved in Evertse and Győry [32], Proposition 3.4 and Lemma 3.2, (i), while (ii) is proved in [32], Lemma 3.6.

□

Now let us describe the above-mentioned representation for the elements of the field K. Recall that D denotes the degree of K over K_0. Since $K = K_0(w)$ for every element $\alpha \in K$ there exists a unique representation $\sum_{j=0}^{D-1} R_{\alpha,j} w^j$, where $R_{\alpha,j} \in K_0$. Since A_0 is

a unique factorization domain (indeed, $z_1, \dots, z_q$ are algebraically independent) and K_0 is its quotient field, thus there exist $P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha \in A_0$ such that

$$\alpha = Q_\alpha^{-1} \sum_{j=0}^{D-1} P_{\alpha,j} w^j \quad \text{with} \quad Q_\alpha \neq 0, \quad \gcd(P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha) = 1. \tag{6.6}$$

Further, the tuple $(P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha)$ is up to sign uniquely determined. Now we define

$$\begin{cases} \overline{\deg}\,\alpha := \max(\deg P_{\alpha,0}, \dots, \deg P_{\alpha,D-1}, \deg Q_\alpha) \\ \overline{h}(\alpha) := \max(h(P_{\alpha,0}), \dots, h(P_{\alpha,D-1}), h(Q_\alpha)), \end{cases} \tag{6.7}$$

if $q > 0$, and $\overline{\deg}\,\alpha = 0$ and $\overline{h}(\alpha) = \log \max(|P_{\alpha,0}|, \dots, |P_{\alpha,D-1}|, |Q_\alpha|)$ if $q = 0$. These two concepts provide a convenient way to measure elements of K.

Proposition 6.2. *Let $\alpha_1, \dots, \alpha_k \in K^*$ and suppose that there are $\tilde{d} > 1$ and $\tilde{h} > 1$ such that $\overline{\deg}\,\alpha_i \leq \tilde{d}$ and $\overline{h}(\alpha_i) \leq \tilde{h}$ for $i = 1, \dots, k$. Then there is a non-zero $f \in A_0$ such that with $B := A_0[w, f^{-1}]$ we have*

$$\begin{aligned} &A \subseteq B, \\ &\alpha_1, \dots, \alpha_k \in B^*. \end{aligned} \tag{6.8}$$

Further, f can be chosen such that it fulfils

$$\begin{aligned} &\deg f \leq (2d_0)^{\exp O(r)} + 2k\tilde{d}, \\ &h(f) \leq (2d_0)^{\exp O(r)}(h_0 + 1) + 2k\tilde{h} + 2rk\tilde{d}. \end{aligned} \tag{6.9}$$

Proof. This proposition is just a variant of (ii) of Propositions 6.1. To prove it we only need to slightly modify the proof of Lemma 3.6. in [32]. In principle, the element f is chosen to be the same as in (ii) of Propositions 6.1 (as constructively given in the proof of Lemma 3.6. in [32]), just the estimate for the degree and height of f is computed in terms of $d_0, \tilde{d}, \tilde{h}$ instead of d^{**} and h^{**}. □

The following two lemmas describe how $\overline{\deg}\,\alpha$ and $\overline{h}(\alpha)$ and the height and degree of representatives for α may be bounded in terms of each other.

Lemma 6.3. *Let $\alpha \in K^*$ and let (a, b) be a pair of representatives for α with $a, b \in \mathbb{Z}[X_1, \dots, X_r]$, $b \notin I$. Put*

$$d^* := \max(d_0, \deg a, \deg b) \quad \textit{and} \quad h^* := \max(h_0, h(a), h(b)).$$

Then

$$\overline{\deg}\,\alpha \leq (2d^*)^{\exp O(r)}, \qquad \overline{h}(\alpha) \leq (2d^*)^{\exp O(r)}(h^* + 1). \tag{6.10}$$

Proof. This is Lemma 3.5 in Evertse and Győry [32]. □

Lemma 6.4. *Let α be a nonzero element of A, and put*

$$\widehat{d} := \max(d_0, \overline{\deg}\,\alpha), \qquad \widehat{h} := \max(h_0, \overline{h}(\alpha)).$$

Then α has a representative $\tilde{\alpha} \in \mathbb{Z}[X_1, \ldots, X_r]$ such that

$$\begin{cases} \deg \tilde{\alpha} \le (2\widehat{d})^{\exp O(r \log^* r)}(\widehat{h}+1), \\ h(\tilde{\alpha}) \le (2\widehat{d})^{\exp O(r \log^* r)}(\widehat{h}+1)^{r+1}. \end{cases} \tag{6.11}$$

Moreover, if $\alpha \in A^*$ then α^{-1} has a representative $\tilde{\alpha}' \in \mathbb{Z}[X_1, \ldots, X_r]$ with

$$\begin{cases} \deg \tilde{\alpha}' \le (2\widehat{d})^{\exp O(r \log^* r)}(\widehat{h}+1), \\ h(\tilde{\alpha}') \le (2\widehat{d})^{\exp O(r \log^* r)}(\widehat{h}+1)^{r+1}. \end{cases} \tag{6.12}$$

Proof. This is a special case of Lemma 3.7 of Evertse and Győry [32] with the choice $\lambda = 1$ and $a = b = 1$. The proof of this lemma is based on work of Aschenbrenner [1]. □

6.2 Using function field results for bounding the degree $\overline{\deg}$ of elements of B

In this section we collect the main tools needed to use results over function fields to bound the $\overline{\deg}$ of elements of B.

We recall some definitions and facts concerning function fields. Let $\Bbbk$ be an algebraically closed field of characteristic 0, z a transcendental element over $\Bbbk$ and M a finite extension of $\Bbbk(z)$. Denote by $\mathcal{M}_M$ the collection of valuations of $M/\Bbbk$; these are the discrete valuations of M with value group $\mathbb{Z}$ that are trivial on $\Bbbk$. Recall that these valuations satisfy the sum formula

$$\sum_{v \in \mathcal{M}_M} v(\alpha) = 0 \qquad \text{for} \quad \alpha \in M^*.$$

The (homogeneous) height of $\mathbf{a} = (\alpha_1, \ldots, \alpha_l) \in M^l$ relative to $M/\Bbbk$ is defined by

$$H_M(\mathbf{a}) = H_M(\alpha_1, \ldots, \alpha_l) := -\sum_{v \in \mathcal{M}_M} \min(v(\alpha_1), \ldots, v(\alpha_l)).$$

The height of $\alpha \in M$ relative to $M/\Bbbk$ is defined by

$$H_M(\alpha) := H_M(1, \alpha) = -\sum_{v \in \mathcal{M}_M} \min(0, v(\alpha)).$$

We will need the following lemma:

Lemma 6.5. *Let $\alpha_1, \dots, \alpha_l \in M$ and suppose that*

$$X^l + f_1 X^{l-1} + \cdots + f_l = (X - \alpha_1) \dots (X - \alpha_l)$$

for certain $f_1, \dots f_l \in \Bbbk[z]$. Then

$$[M : \Bbbk(z)] \max(\deg f_1, \dots, \deg f_l) = \sum_{i=1}^{l} H_M(\alpha_i).$$

Proof. This is Lemma 4.1 in Evertse and Győry [32]. □

Recall that $A = \mathbb{Z}[z_1, \dots, z_r]$ and K denotes the quotient field of A. We keep our assumption that $z_1, \dots, z_q$ is a transcendence basis of $K/\mathbb{Q}$, and the notation $A_0 := \mathbb{Z}[z_1, \dots, z_q]$, $K_0 := \mathbb{Q}(z_1, \dots, z_q)$. Let $w \in A_0$ and $f \in A_0$ be the elements specified in Proposition 6.1 and put $B := A_0[f^{-1}, w]$. Then $K = K_0(w)$ and $A \subseteq B \subset K$. Further, let us denote by $w^{(1)} = w, \dots, w^{(D)}$ the conjugates of w over K_0.

Now we fix $i \in \{1, \dots, q\}$ and introduce the following notation:

$$\begin{aligned} &\Bbbk_i := \mathbb{Q}(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_q), \\ &\overline{\Bbbk}_i \text{ denotes the algebraic closure of } \Bbbk_i, \\ &M_i := \overline{\Bbbk}_i(z_i, w^{(1)}, \dots, w^{(D)}), \\ &B_i := \overline{\Bbbk}_i[z_i, f^{-1}, w^{(1)}, \dots, w^{(D)}]. \end{aligned} \tag{6.13}$$

Clearly, M_i is the splitting field of the minimal polynomial $\mathcal{F}(X)$ of w over the field $\overline{\Bbbk}_i(z_i)$, and B_i is a subring of M_i containing B. Further, we use the following notation:

$$\begin{aligned} &\Delta_i := [M_i : \overline{\Bbbk}_i(z_i)], \\ &g_{M_i} \text{ denotes the genus of } M_i/\overline{\Bbbk}_i, \\ &H_{M_i} \text{ denotes the height with respect to } M_i/\overline{\Bbbk}_i. \end{aligned} \tag{6.14}$$

Put

$$d_1 := \max\{d_0, \deg f, \deg \mathcal{F}_1, \dots, \deg \mathcal{F}_D\}. \tag{6.15}$$

The following Lemma gives an upper bound for the $\overline{\deg}$ of an element of K^* depending on the function field heights with respect to $M_i/\overline{\Bbbk}_i$ of its conjugates.

Lemma 6.6. *Let $\alpha \in K^*$ and denote by $\alpha^{(1)}, \dots, \alpha^{(D)}$ the conjugates of α corresponding to $w^{(1)}, \dots, w^{(D)}$. Then*

$$\overline{\deg}\, \alpha \le qDd_1 + \sum_{i=1}^{q} \Delta_i^{-1} \sum_{j=1}^{D} H_{M_i}(\alpha^{(j)}).$$

Proof. This is Lemma 4.4 in Evertse and Győry [32]. □

Conversely, we have the following:

Lemma 6.7. *Let $\alpha \in K^*$ and $\alpha^{(1)}, \dots, \alpha^{(D)}$ be as in Lemma 6.6. Then we have*

$$\max_{i,j} H_{M_i}(\alpha^{(j)}) \le \Delta_i \left(2D\overline{\deg}\,\alpha + (2d_0)^{\exp O(r)}\right). \tag{6.16}$$

Proof. Consider the representation of the form (6.6) of α. Since $P_{\alpha,k}, Q \in K_0$, we have

$$\alpha^{(j)} = \sum_{k=0}^{D-1} \frac{P_{\alpha,k}}{Q} \left(w^{(j)}\right)^k \qquad \text{for} \quad j = 1, \dots, D.$$

By basic properties of the function field height it follows that

$$H_{M_i}(\alpha^{(j)}) \le \sum_{k=0}^{D-1} H_{M_i}\left(\frac{P_{\alpha,k}}{Q}\right) + \sum_{k=0}^{D-1} k H_{M_i}\left(w^{(j)}\right). \tag{6.17}$$

But we have

$$\begin{aligned} H_{M_i}\left(\frac{P_{\alpha,k}}{Q}\right) &\le \Delta_i H_{\Bbbk_i(z)}\left(\frac{P_{\alpha,k}}{Q}\right) \le \Delta_i(\deg_{z_i} P_{\alpha,k} + \deg_{z_i} Q) \\ &\le \Delta_i(\deg P_{\alpha,k} + \deg Q) \le 2\Delta_i \overline{\deg}\,\alpha. \end{aligned} \tag{6.18}$$

Further, applying Lemma 6.5 with $M_i, w^{(1)}, \dots, w^{(D)}$ instead of $M, \alpha_1, \dots, \alpha_l$, we get

$$\begin{aligned} H_{M_i}\left(w^{(j)}\right) &\le \Delta_i \max_{1\le j\le D}(\deg_{z_i} \mathcal{F}_j) \\ &\le \Delta_i \max_{1\le j\le D}(\deg \mathcal{F}_j) \le \Delta_i (2d_0)^{\exp O(r)}. \end{aligned} \tag{6.19}$$

Now using the fact that $D \le d_0^{r-q} \le d_0^{r-1}$, (6.17), (6.18) and (6.19) imply (6.16). □

6.3 Specializations

Recall again that

$$\begin{aligned} K_0 &= \mathbb{Q}(z_1, \dots, z_q), \quad K = \mathbb{Q}(z_1, \dots, z_q), \\ A_0 &= \mathbb{Z}[z_1, \dots, z_q], \quad B = \mathbb{Z}[z_1, \dots, z_q, w, f^{-1}], \end{aligned} \tag{6.20}$$

where w, f are the elements specified in Proposition 6.1, and the minimal polynomial of w has the form $\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X]$, with degree and coefficients bounded as described in Proposition 6.1.

For every $\mathbf{u} \in \mathbb{Z}^q$ the substitution $z_i \to u_i$ for $i = 1, \dots, q$ defines a mapping from a subring of K_0 to $\overline{\mathbb{Q}}$. More precisely, we fix $\mathbf{u}$ and we consider the ring homomorphism $\varphi_{\mathbf{u}}$ from a subring of K_0 to $\overline{\mathbb{Q}}$ defined by

$$\varphi_{\mathbf{u}}(\alpha) := \alpha(\mathbf{u}) = \frac{g_1(\mathbf{u})}{g_2(\mathbf{u})}$$

for every $\alpha = \frac{g_1}{g_2} \in K_0$ with $g_1, g_2 \in A_0$, and with the additional property $g_2(\mathbf{u}) \neq 0$. To extend this map to a ring homomorphism from B to $\overline{\mathbb{Q}}$ we will impose some restrictions on $\mathbf{u}$. Let $\Delta_{\mathcal{F}}$ denote the discriminant of $\mathcal{F}$ with the convention $\Delta_{\mathcal{F}} = 1$ if $\mathcal{F}$ is a linear polynomial. Put

$$\mathcal{H} := \Delta_{\mathcal{F}} \cdot \mathcal{F}_D \cdot f \in A_0,$$

and assume that $\mathbf{u}$ is chosen such that $\mathcal{H}(\mathbf{u}) \neq 0$. Put

$$\begin{cases} d_0^* = \max(\deg \mathcal{F}_1, \dots, \deg \mathcal{F}_D) \\ h_0^* = \max(h(\mathcal{F}_1), \dots, h(\mathcal{F}_D)) \end{cases} \qquad \begin{cases} d_1^* = \max(d_0^*, \deg f) \\ h_1^* = \max(h_0^*, h(f)). \end{cases} \tag{6.21}$$

Thus we clearly have

$$\deg \mathcal{H} \leq (2D-2) \cdot d_0^* + d_0^* + d_1^* \leq (2D-1) \cdot d_0^* + d_1^*. \tag{6.22}$$

Let us fix a tuple $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$. Since $\mathcal{H}(\mathbf{u}) \neq 0$ implies $\Delta_{\mathcal{F}} \neq 0$ and $\mathcal{F}_D(\mathbf{u}) \neq 0$, the polynomial

$$\mathcal{F}_{\mathbf{u}} := X^D + \mathcal{F}_1(\mathbf{u}) X^{D-1} + \cdots + \mathcal{F}_D(\mathbf{u})$$

has D distinct non-zero roots, say $w^{(1)}(\mathbf{u}), \dots, w^{(D)}(\mathbf{u})$.

Now we extend the map $\varphi_{\mathbf{u}}$ to the ring B in D different ways. Namely, for each $j = 1, \dots, D$ we shall define the function $\varphi_{\mathbf{u},j}$ on B such that if $\alpha \in B$ is written as

$$\alpha = \sum_{i=1}^{D-1} (P_i/Q)\, w^i, \qquad \text{where } P_0, \dots, P_{D-1}, Q \in A_0,\ \gcd(P_0, \dots, P_{D-1}, Q) = 1, \tag{6.23}$$

then

$$\varphi_{\mathbf{u},j}(\alpha) = \alpha^{(j)}(\mathbf{u}) := \sum_{i=1}^{D-1} (P_i(\mathbf{u})/Q(\mathbf{u})) \left(w^{(j)}(\mathbf{u})\right)^i. \tag{6.24}$$

Since $\alpha \in B$, the polynomial Q must divide a power of f and hence $Q(\mathbf{u}) \neq 0$, so $\varphi_{\mathbf{u},j}(\alpha)$ is well-defined. Clearly, $\varphi_{\mathbf{u},j}$ is a ring homomorphism from B to $\overline{\mathbb{Q}}$, thus any of the specializations $\varphi_{\mathbf{u},j}$ maps any unit of B to a non-zero element of $\overline{\mathbb{Q}}$. We mention that $\varphi_{\mathbf{u},j}$ is the identity on $B \cap \mathbb{Q}$. Further, if $\alpha \in B \cap \overline{\mathbb{Q}}$ then $\varphi_{\mathbf{u},j}(\alpha)$ is a conjugate of α

For $\mathbf{u} = (u_1, \ldots, u_q) \in \mathbb{Z}^q$, put $|\mathbf{u}| := \max(|u_1|, \ldots, |u_q|)$. Then for any $g \in A_0$, $\mathbf{u} \in \mathbb{Z}^q$

$$\log |g(\mathbf{u})| \leq q \log \deg g + h(g) + \deg g \log \max(1, |\mathbf{u}|). \tag{6.25}$$

Thus, we have

$$h(\mathcal{F}_{\mathbf{u}}) \leq q \log d_0^* + h_0^* + d_0^* \log \max(1, |\mathbf{u}|) \tag{6.26}$$

and so by Lemma 5.1 of Evertse and Győry [32]

$$\sum_{j=1}^{D} h(w^{(j)}(\mathbf{u})) \leq D + 1 + q \log d_0^* + h_0^* + d_0^* \log \max(1, |\mathbf{u}|). \tag{6.27}$$

Define the algebraic number fields

$$K_{\mathbf{u},j} := \mathbb{Q}(w^{(j)}(\mathbf{u})) \qquad \text{for} \qquad j = 1, \ldots, D, \tag{6.28}$$

and denote by $\Delta_{K_{\mathbf{u},j}}$ their discriminant.

The following lemmas of Evertse and Győry [32] summarize important properties of the above-defined specializations.

Lemma 6.8. *Let $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$. Then for $j = 1, \ldots, D$ we have $[K_{\mathbf{u},j} : \mathbb{Q}] \leq D$ and*

$$|\Delta_{K_{\mathbf{u},j}}| \leq D^{2D-1} \left((d_0^*)^q e^{h_0^*} \max(1, |\mathbf{u}|^{d_0^*}) \right)^{2D-2}.$$

Proof. This is Lemma 5.5 in Evertse and Győry [32]. □

The next lemma bounds the height of $\alpha^{(j)}(\mathbf{u})$ for $\mathbf{u} \in \mathbb{Z}^q$ in terms of the size of $\alpha \in B$ and some parameters of B.

Lemma 6.9. *Let $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$, and let $\alpha \in B$. Then for $j = 1, \ldots, D$,*

$$\begin{aligned} h(\alpha^{(j)}(\mathbf{u})) \leq D^2 + q(D \log d_0^* + \log \overline{\deg}\, \alpha) + \\ + D h_0^* + \overline{h}(\alpha) + (D d_0^* + \overline{\deg}\, \alpha) \log \max(1, |\mathbf{u}|). \end{aligned}$$

Proof. This is Lemma 5.6 in Evertse and Győry [32]. □

The below lemma shows that if we take a sufficiently large number of specializations, then there is at least one specialization among them (say corresponding to $\mathbf{u} \in \mathbb{Z}^q$), such that $\overline{h}(\alpha)$ for $\alpha \in B$ can be bounded by the heights of the images of α by the specializations $\varphi_{\mathbf{u},j}$ for $j = 1, \ldots, D$.

Lemma 6.10. *Let $\alpha \in B$, $\alpha \neq 0$, and let N_0 be an integer with*

$$N_0 \geq \max(\overline{\deg}\,\alpha, 2Dd_0^* + 2(q+1)(d_1^*+1)). \tag{6.29}$$

Then the set

$$\mathcal{S} := \{\mathbf{u} \in \mathbb{Z}^q \ : \ |\mathbf{u}| \leq N_0, \mathcal{H}(\mathbf{u}) \neq 0\}$$

is non-empty, and

$$\overline{h}(\alpha) \leq 5N_0^4(h_1^*+1)^2 + 2D(h_1^*+1)H, \tag{6.30}$$

where $H := \max\{h(\alpha^{(j)}(\mathbf{u})) \ : \ \mathbf{u} \in \mathcal{S}, \ j = 1, \ldots, D\}$.

Proof. This is Lemma 5.7 in Evertse and Győry [32]. □

Chapter 7

Proof of the results from Section 3.2

7.1 A reduction

We shall reduce our equations to equations of the same type over an integral domain $B \supseteq A$ of a special type which is more convenient to deal with.

As before, let $A = \mathbb{Z}[z_1, \ldots, z_r]$ be an integral domain which is finitely generated over $\mathbb{Z}$ and let K be the quotient field of A. Suppose that K has transcendence degree $q \geq 0$. If $q > 0$, we assume without loss of generality that $\{z_1, \ldots, z_q\}$ forms a transcendence basis of $K/\mathbb{Q}$. We define

$$\begin{aligned} &A_0 := \mathbb{Z}[z_1, \ldots, z_q], \qquad && K_0 := \mathbb{Q}(z_1, \ldots, z_q) && \text{if } q > 0 \\ &A_0 := \mathbb{Z}, && K_0 := \mathbb{Q} && \text{if } q = 0. \end{aligned}$$

The field K is a finite extension of K_0. Further, if $q = 0$, it is an algebraic number field. In case that $q > 0$, for $f \in A_0 \setminus \{0\}$ we define $\deg f$ and $h(f)$ to be the total degree and logarithmic height of f, viewed as a polynomial in the variables $z_1, \ldots, z_q$. In case that $q = 0$, for $f \in A_0 \setminus \{0\} = \mathbb{Z} \setminus \{0\}$, we put $\deg f := 0$ and $h(f) := \log |f|$.

We shall construct an integral extension B of A in K such that

$$B := A_0[w, g^{-1}], \tag{7.1}$$

where $g \in A_0 \setminus \{0\}$ and w is a primitive element of K over K_0 that is integral over A_0. Then we give a bound for the sizes of the solutions of our equations in $x, y \in B$.

We recall that $A \cong \mathbb{Z}[X_1, \ldots, X_r]/\mathcal{I}$ where $\mathcal{I} \subset \mathbb{Z}[X_1, \ldots, X_r]$ is the ideal of polynomials f with $f(z_1, \ldots, z_r) = 0$ and z_i corresponds to the residue class of X_i modulo $\mathcal{I}$. The ideal $\mathcal{I}$ is finitely generated. Assume that

$$\mathcal{I} = (f_1, \ldots, f_t),$$

and put

$$d_0 := \max(1, \deg f_1, \ldots, \deg f_t), \qquad h_0 := \max(1, h(f_1), \ldots, h(f_t)). \tag{7.2}$$

We shall use Proposition 6.1 (ii) in a special case. To state it, we introduce some further notation and prove a lemma.

We recall that $a_0, a_1, \ldots, a_n \in A$ are the coefficients of the binary form $F(X, Y)$, resp. of the polynomial $F(X)$ on the right hand side of our Thue equation, resp. of our hyper- or superelliptic equation, and $\tilde{a}_0, \tilde{a}_1, \ldots, \tilde{a}_n$ denote their representatives satisfying (3.4) resp. (3.7). This implies that $d_0 \leq d$, $h_0 \leq h$, and that $\tilde{a}_i$ has total degree $\leq d$ and logarithmic height $\leq h$ for $i = 0, \ldots, n$. Denote by $\tilde{F}$ the binary form $F(X, Y)$ resp. the polynomial $F(X)$ with coefficients $a_0, a_1, \ldots, a_n$ replaced by $\tilde{a}_0, \tilde{a}_1, \ldots, \tilde{a}_n$, and by $D_{\tilde{F}}$ the discriminant of $\tilde{F}$. In view of the assumption $D_F \neq 0$ we have $D_{\tilde{F}} \notin \mathcal{I}$.

Keeping the notation and assumptions of Section 3.2 we have the following lemma.

Lemma 7.1. *For the discriminant $D_{\tilde{F}}$ the following statements are true:*

$$\deg D_{\tilde{F}} \leq (2n-2)d, \tag{7.3}$$

$$h(D_{\tilde{F}}) \leq (2n-2)\left(\log\left(2n^2\binom{d+r}{r}\right) + h\right). \tag{7.4}$$

Proof. Recall that the discriminant $D_{\tilde{F}}$ can be expressed as

$$D(\tilde{F}) = \pm \begin{vmatrix} \tilde{a}_0 & \tilde{a}_1 & \cdots & \cdots & \tilde{a}_n & & \\ & \ddots & & & & \ddots & \\ & & \tilde{a}_0 & \tilde{a}_1 & \cdots & \cdots & \tilde{a}_n \\ \tilde{a}_1 & 2\tilde{a}_2 & \cdots & n\tilde{a}_n & & & \\ n\tilde{a}_0 & (n-1)\tilde{a}_1 & \cdots & \tilde{a}_{n-1} & & & \\ & \ddots & & & \ddots & & \\ & & \ddots & & & \ddots & \\ & & & n\tilde{a}_0 & (n-1)\tilde{a}_1 & \cdots & \tilde{a}_{n-1} \end{vmatrix}, \tag{7.5}$$

with on the first $n-2$ rows of the determinant $\tilde{a_0}, \ldots, \tilde{a_n}$, then on the $(n-1)$-st row $\tilde{a}_1, 2\tilde{a}_2, \ldots, n\tilde{a}_n$, and on the last $n-1$ rows $n\tilde{a}_0, \ldots, \tilde{a}_{n-1}$. This implies at once (7.3).

To prove (7.4), we use the length $L(P)$ of a polynomial $P \in \mathbb{Z}[X_1, \ldots, X_r]$, that is the sum of the absolute values of the coefficients of P. It is known and easily seen that if $P, Q \in \mathbb{Z}[X_1, \ldots, X_r]$ then $L(P+Q)$ and $L(PQ)$ do not exceed $L(P)+L(Q)$ and $L(P)L(Q)$, respectively (see e.g. Waldschmidt [77], p.76).

We have

$$L(\tilde{a}_i) \le \binom{d+r}{r} H \qquad \text{with} \quad H = \exp h \qquad \text{for } i = 0, \dots, n.$$

By applying these facts to (7.5), we obtain

$$L(D_{\tilde{F}}) \le (2n-2)! \left(n \binom{d+r}{r} H \right)^{2n-2}.$$

Together with $h(D_{\tilde{F}}) \le \log L(D_{\tilde{F}})$ this implies (7.4). □

We now apply Proposition 6.1 (ii) to the numbers $\alpha_1 = \delta$, $\alpha_2 = \delta^{-1}$, $\alpha_3 = D_F$ and $\alpha_4 = D_F^{-1}$. Then the pairs $(\tilde{\delta}, 1)$, $(1, \tilde{\delta})$, $(D_{\tilde{F}}, 1)$, $(1, D_{\tilde{F}})$ represent the numbers α_i, $i = 1, \dots, 4$. Using the upper bounds for $\deg D_{\tilde{F}}$, $h(D_{\tilde{F}})$ implied by Lemma 7.1 as well as the upper bounds $\deg \tilde{\delta} \le d$, $h(\tilde{\delta}) \le h$ implied by (3.4), (3.7), we get immediately from Proposition 6.1 (ii) the following.

Proposition 7.2. *There is a non-zero $g \in A_0$ such that*

$$A \subseteq A_0[w, g^{-1}], \quad \delta, D_F \in A_0[w, g^{-1}]^* \tag{7.6}$$

and

$$\deg g \le (nd)^{\exp O(r)}, \qquad h(g) \le (nd)^{\exp O(r)}(h+1). \tag{7.7}$$

We recall some notation introduced in Chapter 6. In the case $q > 0$, $z_1, \dots, z_q$ are algebraically independent. Thus, for $q \ge 0$, A_0 is a unique factorization domain, and hence the greatest common divisor of a finite set of elements of A_0 is well defined and up to sign uniquely determined. We associate with every element $\alpha \in K$ the up to sign unique tuple $P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha$ of elements of A_0 such that

$$\alpha = Q_\alpha^{-1} \sum_{j=0}^{D-1} P_{\alpha,j} w^j \quad \text{with} \quad Q_\alpha \ne 0, \quad \gcd(P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha) = 1. \tag{7.8}$$

We put

$$\begin{cases} \overline{\deg}\, \alpha := \max(\deg P_{\alpha,0}, \dots, \deg P_{\alpha,D-1}, \deg Q_\alpha) \\ \overline{h}(\alpha) := \max(h(P_{\alpha,0}), \dots, h(P_{\alpha,D-1}), h(Q_\alpha)), \end{cases} \tag{7.9}$$

where as usual, $\deg P$, $h(P)$ denote the total degree and logarithmic height of a polynomial P with rational integral coefficients. Thus for $q = 0$ we have $\overline{\deg}\, \alpha = 0$ and $\overline{h}(\alpha) = \log \max(|P_{\alpha,0}|, \dots, |P_{\alpha,D-1}|, |Q_\alpha|)$. We mention that if $\alpha \in K^*$ has a representation pair (a, b) then $\overline{\deg}\, \alpha$ and $\overline{h}(\alpha)$ can be estimated in terms of the degrees and heights of a and b (see Lemma 6.3) and every $\alpha \in A$ has a representative whose degree and height can be bounded above in terms of $\overline{\deg}\, \alpha$ and $\overline{h}(\alpha)$ (see Lemma 6.4).

7.1.1 Thue equations

Recall that $A_0 = \mathbb{Z}[z_1, \dots, z_q]$, $K_0 = \mathbb{Q}(z_1, \dots, z_q)$ if $q > 0$, and $A_0 = \mathbb{Z}$, $K_0 = \mathbb{Q}$ if $q = 0$, and that in the case $q = 0$ total degrees and $\overline{\deg}$-s are always zero. Further, we have

$$F(X, Y) = a_0 X^n + a_1 X^{n-1} Y + \cdots + a_n Y^n \in A[X, Y]$$

with $n \geq 3$ and with discriminant $D_F \neq 0$, and $\delta \in A \setminus \{0\}$. Recall that for $a_0, a_1, \dots, a_n, \delta$ we have chosen representatives $\tilde{a_0}, \tilde{a_1}, \dots, \tilde{a_n}, \tilde{\delta} \in \mathbb{Z}[X_1, \dots, X_r]$ satisfying (3.4).

Theorem 3.1 will be deduced from the following Proposition, which makes sense also if $q = 0$. The proof of this proposition is given in Sections 7.2–7.4.

Proposition 7.3. *Let w and g be as in Propositions 6.1 (i) and 7.2, respectively, with the properties specified there, and consider the integral domain*

$$B := A_0[w, g^{-1}].$$

Then for the solutions x, y of the equation

$$F(x, y) = \delta \qquad \text{in} \quad x, y \in B \tag{7.10}$$

we have

$$\overline{\deg}\, x,\ \overline{\deg}\, y \leq (nd)^{\exp O(r)}, \tag{7.11}$$

$$\overline{h}(x),\ \overline{h}(y) \leq \exp\left(n!(nd)^{\exp O(r)}(h+1)\right). \tag{7.12}$$

We now deduce Theorem 3.1 from Proposition 7.3.

Proof of Theorem 3.1. Let x, y be a solution of equation (3.3). In view of (7.6) x, y is also a solution in $B = A_0[w, g^{-1}]$, where g, w satisfy the properties specified in Propositions 6.1 (i) and 7.2, respectively. Then by Proposition 7.3, the inequalities (7.11) and (7.12) hold. Applying now Lemma 6.4 to x and y, we infer that x, y have representatives $\tilde{x}, \tilde{y}$ in $Z[X_1, \dots, X_r]$ with (3.5). □

7.1.2 Hyper- and superelliptic equations

Recall that the polynomial

$$F(X) = a_0 X^n + a_1 X^{n-1} + \cdots + a_n \in A[X]$$

has discriminant $D_F \neq 0$, that $\delta \in A \setminus \{0\}$, and that for $a_0, a_1, \dots, a_n, \delta$ we have chosen representatives $\tilde{a_0}, \tilde{a_1}, \dots, \tilde{a_n}, \tilde{\delta} \in \mathbb{Z}[X_1, \dots, X_r]$ satisfying (3.7).

Theorem 3.3 will be deduced from the following Proposition, which has a meaning also if $q = 0$. Similarly as its analogue for Thue equations, its proof is given in Sections 7.2–7.4.

Proposition 7.4. *Let w and g be as in Propositions 6.1 (i) and 7.2, respectively, with the properties specified there, and consider the domain*

$$B := A_0[w, g^{-1}].$$

Further, let m be an integer ≥ 2, and assume that $n \geq 3$ if $m = 2$ and $n \geq 2$ if $m \geq 3$. Then for the solutions x, y of the equation

$$F(x) = \delta y^m \qquad \text{in} \quad x, y \in B \tag{7.13}$$

we have

$$\overline{\deg}\, x,\ m\overline{\deg}\, y \leq (nd)^{\exp O(r)}, \tag{7.14}$$

$$\overline{h}(x),\ \overline{h}(y) \leq \exp\left(m^3(nd)^{\exp O(r)}(h+1)\right) \tag{7.15}$$

We now deduce Theorem 3.3 from Proposition 7.4.

Proof of Theorem 3.3. Let x, y be a solution of equation (3.6). In view of (7.6) x, y is also a solution in $B = A_0[w, g^{-1}]$, where g, w satisfy the conditions specified in Propositions 6.1 (i) and 7.2, respectively. Then by Proposition 7.4, the inequalities (7.14) and (7.15) hold. Applying now Lemma 6.4 to x and y, we infer that x, y have representatives $\tilde{x}, \tilde{y}$ in $Z[X_1, \ldots, X_r]$ with (3.8). □

Proposition 7.5. *Suppose that equation (7.13) has a solution $x \in B$, $y \in B \cap \overline{\mathbb{Q}}$ and that also $y \neq 0$ and y is not a root of unity. Then*

$$m \leq \exp\left((nd)^{\exp O(r)}(h+1)\right). \tag{7.16}$$

Proof of Theorem 3.4. Let $x, y \in A$, $m \in \mathbb{Z}_{\geq 2}$ be a solution of equation (3.6). First let $y \notin \overline{\mathbb{Q}}$. Then $\overline{\deg}\, y \geq 1$, and together with (7.14) this implies (3.10). Next, let $y \in \overline{\mathbb{Q}}$. Then Proposition 7.5 gives at once (3.9). □

The proof of Proposition 7.5 is a combination of results from Sections 7.2–7.4. It is completed at the end of Section 7.4.

7.2 Bounding the degree

In this section we shall prove (7.11) of Proposition 7.3 and (7.14) of Proposition 7.4.

We recall some results on function fields in one variable. Let $\Bbbk$ be an algebraically closed field of characteristic 0, z a transcendental element over $\Bbbk$ and M a finite extension of $\Bbbk(z)$. Denote by $g_{M/\Bbbk}$ the genus of M, and by $\mathcal{M}_M$ the collection of valuations of $M/\Bbbk$, these are the discrete valuations of M with value group $\mathbb{Z}$ that are trivial on $\Bbbk$. Recall that these valuations satisfy the sum formula

$$\sum_{v\in\mathcal{M}_M} v(\alpha) = 0 \qquad \text{for} \quad \alpha \in M^*.$$

For a finite subset S of $\mathcal{M}_M$, an element $\alpha \in M$ is called an S-integer if $v(\alpha) \geq 0$ for all $v \in \mathcal{M}_M \setminus S$. The S-integers form a ring in M, denoted by $\mathcal{O}_S$. The (homogeneous) height of $\mathbf{a} = (\alpha_1, \ldots, \alpha_l) \in M^l$ relative to $M/\Bbbk$ is defined by

$$H_M(\mathbf{a}) = H_M(\alpha_1, \ldots, \alpha_l) := -\sum_{v\in\mathcal{M}_M} \min(v(\alpha_1), \ldots, v(\alpha_l)),$$

and we define the height $H_M(f)$ of a polynomial $f \in M[X]$ by the height of the vector defined by the coefficients of f. Further, we shall write $H_M(1, \mathbf{a}) := H_M(1, \alpha_1, \ldots, \alpha_l)$. We note that

$$H_M(\alpha_i) \leq H_M(\mathbf{a}) \leq H_M(\alpha_1) + \cdots + H_M(\alpha_l), \qquad i = 1, \ldots, l. \tag{7.17}$$

By the sum formula,

$$H_M(\alpha\mathbf{a}) = H_M(\mathbf{a}) \qquad \text{for} \quad \alpha \in M^*. \tag{7.18}$$

The height of $\alpha \in M$ relative to $M/\Bbbk$ is defined by

$$H_M(\alpha) := H_M(1, \alpha) = -\sum_{v\in\mathcal{M}_M} \min(0, v(\alpha)).$$

It is clear that $H_M(\alpha) = 0$ if and only if $\alpha \in \Bbbk$. Using the sum formula, it is easy to prove that the height has the properties

$$\begin{aligned} &H_M(\alpha^l) = |l| H_M(\alpha), \\ &H_M(\alpha+\beta) \leq H_M(\alpha) + H_M(\beta), \qquad H_M(\alpha\beta) \leq H_M(\alpha) + H_M(\beta) \end{aligned} \tag{7.19}$$

for all non-zero $\alpha, \beta \in M$ and for every integer l.

If L is a finite extension of M, we have

$$H_L(\alpha_0, \ldots, \alpha_l) = [L : M] H_M(\alpha_0, \ldots, \alpha_l) \qquad \text{for} \quad \alpha_0, \ldots, \alpha_l \in M. \tag{7.20}$$

By $\deg f$ we denote the total degree of $f \in \Bbbk[z]$. For $f_0, \ldots, f_l \in \Bbbk[z]$ with $\gcd(f_0, \ldots, f_l) = 1$ we have

$$H_{\Bbbk[z]}(f_0, \ldots, f_l) = \max(\deg f_0, \ldots, \deg f_l). \tag{7.21}$$

Lemma 7.6. *Let*

$$F = f_0X^l + f_1X^{l-1} + \cdots + f_l \in M[X]$$

be a polynomial with $f_0 \neq 0$ and with non-zero discriminant. Let L be the splitting field over M of F. Then

$$g_{L/\Bbbk} \leq [L : M] \cdot \big(g_{M/\Bbbk} + lH_M(F)\big).$$

In particular, if $M = \Bbbk(z)$ and $f_0, \ldots, f_l \in \Bbbk[z]$, we have

$$g_{L/\Bbbk} \leq [L : M] \cdot l \max(\deg f_0, \ldots, \deg f_l).$$

Proof. The second assertion follows by combining the first assertion with (7.21). We now prove the first assertion. Our proof is a generalization of that of Lemma H of Schmidt [67].

For $v \in \mathcal{M}_M$, put $v(F) := \min(v(f_0), \ldots, v(f_l))$. Let D_F denote the discriminant of F. Since D_F is a homogeneous polynomial of degree $2l-2$ in $f_0, \ldots, f_l$, we have

$$v(D_F) \geq (2l-2)v(F). \tag{7.22}$$

Let S be the set of $v \in \mathcal{M}_M$ with $v(f_0) > v(F)$ or $v(D_F) > (2l-2)v(F)$. We show that L/M is unramified over every valuation $v \in \mathcal{M}_M \setminus S$.

Take $v \in \mathcal{M}_M \setminus S$. Let

$$O_v := \{x \in M : v(x) \geq 0\}, \quad m_v := \{x \in M : v(x) > 0\}$$

denote the local ring at v, and the maximal ideal of O_v, respectively. The residue class field O_v/m_v is equal to $\Bbbk$ since $\Bbbk$ is algebraically closed. Let $\varphi_v : O_v \to \Bbbk$ denote the canonical homomorphism.

Without loss of generality, we assume $v(F) = 0$. Then $v(f_0) = 0$, $v(D_F) = 0$. Let $\varphi_v(F) := \sum_{j=0}^{l} \varphi_v(f_j)X^{l-j}$. Then $\varphi_v(f_0) \neq 0$ and $\varphi_v(F)$ has discriminant $\varphi_v(D_F) \neq 0$. Since $D_F \neq 0$, the polynomial F has l distinct zeros in L, $\alpha_1, \ldots, \alpha_l$, say. Further, $\varphi_v(F)$ has l distinct zeros in $\Bbbk$, $a_1, \ldots, a_l$, say.

Denote by Σ_l the permutation group on $(1, \ldots, l)$. Choose $c_1, \ldots, c_l \in \Bbbk$, such that the numbers

$$\alpha_\sigma := c_1\alpha_{\sigma(1)} + \cdots + c_l\alpha_{\sigma(l)} \quad (\sigma \in \Sigma_l)$$

are all distinct, and the numbers

$$a_\sigma := c_1a_{\sigma(1)} + \cdots + c_la_{\sigma(l)} \quad (\sigma \in \Sigma_l)$$

are all distinct. Let $\alpha := c_1\alpha_1 + \cdots + c_l\alpha_l$. Then $L = M(\alpha)$, and the monic minimal polynomial of α over M divides $G := \prod_{\sigma\in\Sigma_l}(X - \alpha_\sigma)$ which by the theorem of symmetric

functions belongs to $M[X]$. The image of G under φ_v is $\prod_{\sigma\in\Sigma_l}(X-a_\sigma)$ and this has only simple zeros. This implies that L/M is unramified at v.

For $v\in\mathcal{M}_M$ and any valuation $\in\mathcal{M}_L$ above v, denote by $e(V|v)$ the ramification index of V over v. Recall that $\sum_{V|v}e(V|v)=[L:M]$, where the sum is taken over all valuations of L lying above v. Now the Riemann-Hurwitz formula implies that

$$\begin{aligned} 2g_{L/\Bbbk}-2 &= [L:M](2g_{M/\Bbbk}-2)+\sum_{v\in S}\sum_{V|v}(e(V|v)-1) \qquad (7.23)\\ &\le [L:M](2g_{M/\Bbbk}-2+|S|), \end{aligned}$$

where $|S|$ denotes the cardinality of S. It remains to estimate $|S|$. By the sum formula and (7.22) we have

$$\begin{aligned} |S| &\le \sum_{v\in S}\Big((v(f_0)-v(F))+(v(D_F)-(2l-2)v(F))\Big)\\ &= -\sum_{v\in S}(2l-1)v(F)-\sum_{v\in\mathcal{M}_M\setminus S}v(f_0)-\sum_{v\in\mathcal{M}_M\setminus S}v(D_F)\\ &\le -(2l-1)\sum_{v\in\mathcal{M}_M}v(F)=(2l-1)H_M(F). \end{aligned}$$

By inserting this into (7.23) we arrive at an inequality which is stronger than what we wanted to prove. □

In the sequel we keep the notation of Proposition 6.1. To prove (7.11) and (7.14) we may suppose that $q>0$ since the case $q=0$ is trivial. Let again $K_0:=\mathbb{Q}(z_1,\dots,z_q)$, $K:=K_0(w)$, $A_0:=\mathbb{Z}[z_1,\dots,z_q]$, $B:=\mathbb{Z}[z_1,\dots,z_q,w,g^{-1}]$ with g,w specified in Propositions 6.1 (i) and 7.2.

Fix $i\in\{1,\dots,q\}$. Let $\Bbbk_i:=\mathbb{Q}(z_1,\dots,z_{i-1},z_{i+1},\dots,z_q)$ and $\overline{\Bbbk}_i$ its algebraic closure. Then A_0 is contained in $\overline{\Bbbk}_i[z_i]$. Denote by $w^{(1)}:=w,\dots,w^{(D)}$ the conjugates of w over K_0. Let M_i denote the splitting field of the polynomial $X^D+\mathcal{F}_1X^{D-1}+\cdots+\mathcal{F}_D$ over $\overline{\Bbbk}_i(z_i)$, that is

$$M_i:=\overline{\Bbbk}_i(z_i,w^{(1)},\dots,w^{(D)}).$$

Then

$$B_i:=\overline{\Bbbk}_i[z_i,g^{-1},w^{(1)},\dots,w^{(D)}]$$

is a subring of M_i which contains $B=\mathbb{Z}[z_1,\dots,z_q,w,g^{-1}]$ as a subring. Let $\Delta_i:=[M_i:\overline{\Bbbk}_i(z_i)]$. Further, let g_{M_i} denote the genus of $M_i/\overline{\Bbbk}_i$, and H_{M_i} the height taken with respect to $M_i/\overline{\Bbbk}_i$. Put

$$d_1:=\max(d_0,\deg f,\deg\mathcal{F}_1,\dots,\deg\mathcal{F}_D). \qquad (7.24)$$

We mention that in view of Propositions 6.1, 7.2,

$$d_1 \leq (nd)^{\exp O(r)}. \tag{7.25}$$

7.2.1 Thue equations

As before, $\mathbb{k}$ is an algebraically closed field of characteristic 0, z a transcendental element over $\mathbb{k}$ and M a finite extension of $\mathbb{k}(z)$. Further, $g_{M/\mathbb{k}}$ denotes the genus of M, $\mathcal{M}_M$ the collection of valuations of $M/\mathbb{k}$, and for a finite subset S of $\mathcal{M}_M$, $\mathcal{O}_S$ denotes the ring of S-integers in M. We denote by $|S|$ the cardinality of S.

Consider now the Thue equation

$$F(x, y) = 1 \qquad \text{in} \quad x, y \in \mathcal{O}_S, \tag{7.26}$$

where F is a binary form of degree $n \geq 3$ with coefficients in M and with non-zero discriminant.

Proposition 7.7. *Every solution $x, y \in \mathcal{O}_S$ of (7.26) satisfies*

$$\max(H_M(x), H_M(y)) \leq 89 H_M(F) + 212 g_{M/\mathbb{k}} + |S| - 1. \tag{7.27}$$

Proof. This is Theorem 1 (ii) of Schmidt [67]. □

We note that from Mason's fundamental inequality concerning S-unit equations over function fields (see Mason [55]) one could deduce (7.27) with smaller constants than 89 and 212. However, this is irrelevant for the bounds in (3.5).

Now we use Proposition 7.7 to prove the statement (7.11) of Proposition 7.3.

Proof of (7.11). Recall that $w^{(1)} := w, \ldots, w^{(D)}$ are the conjugates of w over K_0, and for $\alpha \in K$ we denote by $\alpha^{(1)}, \ldots, \alpha^{(D)}$ the conjugates of α corresponding to $w^{(1)}, \ldots, w^{(D)}$.

Recall also that for $i = 1, \ldots, q$ we defined $\mathbb{k}_i := \mathbb{Q}(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_q)$ and $\overline{\mathbb{k}}_i$ denotes its algebraic closure. Further, M_i denotes the splitting field of the polynomial $X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D$ over $\overline{\mathbb{k}}_i(z_i)$. We put $\Delta_i := [M_i : \overline{\mathbb{k}}_i(z_i)]$ and define

$$S_i := \{v \in \mathcal{M}_{M_i} \ : \ v(z_i) < 0 \text{ or } v(g) > 0\}.$$

The conjugates $w^{(j)}$ $(j = 1, \ldots, D)$ lie in M_i and are all integral over $\mathbb{k}_i[z_i]$. Hence they belong to $\mathcal{O}_{S_i}$. Further, $g^{-1} \in \mathcal{O}_{S_i}$. Consequently, if $\alpha \in B = A_0[w, g^{-1}]$, then $\alpha^{(j)} \in \mathcal{O}_{S_i}$ for $j = 1, \ldots, D$, $i = 1, \ldots, q$.

Let x, y be a solution of equation (7.10). Put $F' := \delta^{-1}F$, and let $F'^{(j)}$ be the binary form obtained by taking the j-th conjugates of the coefficients of F'. Let $j \in \{1, \dots, D\}$, $i \in \{1, \dots, q\}$. Then clearly, $F'^{(j)} \in M_i[X, Y]$, and

$$F'^{(j)}(x^{(j)}, y^{(j)}) = 1, \qquad x^{(j)}, y^{(j)} \in \mathcal{O}_{S_i}.$$

So by Proposition 7.7 we obtain that

$$\max(H_{M_i}(x^{(j)}), H_{M_i}(y^{(j)})) \le 89 H_{M_i}(F^{(j)}) + 212 g_{M_i} + |S_i| - 1. \tag{7.28}$$

We estimate the various parameters in this bound. We start with $H_{M_i}(F'^{(j)})$. We recall that $F'(X, Y) = \delta^{-1}(a_0 X^n + a_1 X^{n-1}Y + \cdots + a_n Y^n)$. Using (7.18), (7.17) and Lemma 6.7 we infer that

$$\begin{aligned} H_{M_i}(F'^{(j)}) &= H_{M_i}(a_0^{(j)}, \dots, a_n^{(j)}) \le H_{M_i}(a_0^{(j)}) + \cdots + H_{M_i}(a_n^{(j)}) \\ &\le \Delta_i \left(2D(\overline{\deg}\, a_0 + \cdots + \overline{\deg}\, a_n) + n(2d_0)^{\exp O(r)}\right). \end{aligned}$$

By Lemma 6.3 we have

$$\overline{\deg}\, a_i \le (2d^*)^{\exp O(r)} \qquad \text{for} \quad i = 0, \dots, n,$$

where $d^* := \max(d_0, \deg \tilde{a}_i) \le d$. Further, we have $d_0 \le d$, $D \le d_0^{r-q} \le d^r$. Thus we obtain that

$$\begin{aligned} H_{M_i}(F'^{(j)}) &\le \Delta_i \big(2D(n+1)(2d)^{\exp O(r)} + n(2d)^{\exp O(r)}\big) \\ &\le \Delta_i (nd)^{\exp O(r)}. \end{aligned} \tag{7.29}$$

Next, we estimate the genus. Using Lemma 7.6 with $F(X) = \mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D$, applying Proposition 6.1, and using $d_0 \le d$, $D \le d_0^r \le d^r$, we infer that

$$g_{M_i} \le \Delta_i D \max_{1 \le k \le D} \deg_{z_i} \mathcal{F}_k \le \Delta_i D (2d_0)^{\exp O(r)} \le \Delta_i (nd)^{\exp O(r)}. \tag{7.30}$$

Lastly, we estimate $|S_i|$. Each valuation of $\overline{\Bbbk}_i(z_i)$ can be extended to at most $[M_i : \overline{\Bbbk}_i(z_i)] = \Delta_i$ valuations of M_i. Thus M_i has at most Δ_i valuations v with $v(z_i) < 0$ and at most $\Delta_i \deg f$ valuations v with $v(f) > 0$. Hence using Proposition 7.2, we get

$$|S_i| \le \Delta_i + \Delta_i \deg_{z_i} f \le \Delta_i (1 + \deg f) \le \Delta_i (nd)^{\exp O(r)}. \tag{7.31}$$

By inserting the bounds (7.29), (7.30) and (7.31) into (7.28), we infer

$$\max(H_{M_i}(x^{(j)}), H_{M_i}(y^{(j)})) \le \Delta_i (nd)^{\exp O(r)}. \tag{7.32}$$

In view of Lemma 6.6, (7.32), $D \le d^r$, $q \le r$ and (7.25) we deduce that

$$\overline{\deg}\, x, \overline{\deg}\, y \le qDd_1 + \sum_{i=1}^{q} \Delta_i^{-1} \sum_{j=1}^{D} H_{M_i}(x^{(j)}) \le (nd)^{\exp O(r)}.$$

This proves (7.11). □

7.2.2 Hyper- and superelliptic equations

Recall the notation introduced at the beginning of Section 7.2. Again, $\Bbbk$ is an algebraically closed field of characteristic 0, z a transcendental element over $\Bbbk$, M a finite extension of $\Bbbk(z)$, and S a finite subset of $\mathcal{M}_M$.

Proposition 7.8. *Let $F \in M[X]$ be a polynomial with non-zero discriminant and $m \geq 3$ a given integer. Put $n := \deg F$ and assume $n \geq 2$. All solutions of the equation*

$$F(x) = y^m \qquad \text{in} \quad x, y \in \mathcal{O}_S \tag{7.33}$$

have the property

$$\begin{aligned} H_M(x) &\leq (6n+18)H_M(F) + 6g_{M/\Bbbk} + 2|S|, && (7.34)\\ mH_M(y) &\leq (6n^2+18n+1)H_M(F) + 6ng_{M/\Bbbk} + 2n|S|. && (7.35) \end{aligned}$$

Proof. First assume that F splits into linear factors over M, and that S consists only of the infinite valuations of M, these are the valuations of M with $v(z) < 0$. Under these hypotheses, Mason [55, p.118, Theorem 15], proved that for every solution x, y of (7.33) we have

$$H_M(x) \leq 18H_M(F) + 6g_{M/\Bbbk} + 2(|S| - 1). \tag{7.36}$$

But Mason's proof remains valid without any changes for any arbitrary finite set of places S. That is, (7.36) holds if F splits into linear factors over M, without any condition on S.

We reduce the general case, where the splitting field of M may be larger than M, to the case considered by Mason. Let L be the splitting field of F over M, and T the set of valuations of L that extend those of S. Then $|T| \leq [L : M] \cdot |S|$, and by Lemma 7.6, we have $g_{L/\Bbbk} \leq [L : M] \cdot (g_{M/\Bbbk} + nH_M(F))$. Note that (7.36) holds, but with L, T instead of M, S. It follows that

$$\begin{aligned} [L : M] \cdot H_M(x) = H_L(x) &\leq 18H_L(F) + 6g_{L/\Bbbk} + 2(|T| - 1)\\ &\leq [L : M]\big((6n+18)H_M(F) + 6g_{M/\Bbbk} + 2|S|\big) \end{aligned}$$

which implies (7.34). Further,

$$mH_M(y) = H_M(y^m) = H_M(F(x)) \leq H_M(F) + nH_M(x), \tag{7.37}$$

which gives (7.35). □

Proposition 7.9. *Let $F \in M[X]$ be a polynomial with non-zero discriminant. Put $n := \deg F$ and assume $n \geq 3$. Then the solutions of*

$$F(x) = y^2 \qquad in \quad x, y \in \mathcal{O}_S \tag{7.38}$$

have the property

$$H_M(x) \leq (42n + 37)H_M(F) + 8g_{M/\Bbbk} + 4|S|, \tag{7.39}$$

$$H_M(y) \leq (21n^2 + 19n)H_M(F) + 4ng_{M/\Bbbk} + 2n|S|. \tag{7.40}$$

Proof. First assume that F splits into linear factors over M, that S consists only of the infinite valuations of M, that F is monic, and that F has its coefficients in $\mathcal{O}_S$. Under these hypotheses, Mason [55, p.30, Theorem 6] proved that for every solution of (7.38) we have

$$H_M(x) \leq 26H_M(F) + 8g_{M/\Bbbk} + 4(|S| - 1). \tag{7.41}$$

An inspection of Mason's proof shows that his result is valid for arbitrary finite sets of valuations S, not just the set of infinite valuations. This leaves only the conditions imposed on F.

We reduce the general case to the special case to which (7.41) is applicable. Let $F = a_0X^n + \cdots + a_n$. Let L be the splitting field of $F \cdot (X^2 - a_0)$ over M. Let T be the set of valuations of L that extend the valuations of S, and also the valuations $v \in \mathcal{M}_M$ such that $v(F) < 0$. Further, let $F' = X^n + a_1X^{n-1} + a_0a_1X^{n-2} + \cdots + a_0^{n-1}a_n$, and let b be such that $b^2 = a_0^{n-1}$. Then for every solution x, y of (7.38) we have

$$F'(a_0x) = (by)^2, \quad a_0x, by \in \mathcal{O}_T,$$

and moreover, $F' \in \mathcal{O}_T[X]$, F' is monic, and F' splits into linear factors over L. So by (7.41),

$$H_L(a_0x) \leq 26H_L(F') + 8g_{L/\Bbbk} + 4(|T| - 1). \tag{7.42}$$

First notice that

$$H_L(F') = [L : M]H_M(F') \leq [L : M] \cdot nH_M(F).$$

Further,

$$|T| \leq [L : M]\Big(|S| - \sum_{v \in \mathcal{M}_M} \min(0, v(F))\Big) \leq [L : M]\big(|S| + H_M(F)\big).$$

Finally, by $H_M(F \cdot (X^2 - a_0)) \leq 2H_M(F)$ and Lemma 7.6, we have

$$g_{L/\Bbbk} \leq [L : M](g_{M/\Bbbk} + (n + 2)2H_M(F)).$$

By inserting these bounds into (7.42), we infer

$$\begin{aligned}[L:M]H_M(x) &\le [L:M]\big(H_M(a_0x)+H_M(F)\big) = H_L(a_0x)+[L:M]H_M(F)\\ &\le [L:M]\big((42n+37)H_M(F)+8g_{M/\Bbbk}+4|S|\big).\end{aligned}$$

This implies (7.39). The other inequality (7.40) follows by combining (7.39) with (7.37) with $m=2$. □

The final step of this subsection is to prove statement (7.14) in Proposition 7.4.

Proof of (7.14). We closely follow the proof of statement (7.11) in Proposition 7.3, and use the same notation. In particular, $\Bbbk_i, M_i, S_i, \Delta_i$ will have the same meaning, and for $\alpha \in B$, $j=1,\dots,D$, the j-th conjugate $\alpha^{(j)}$ is the one corresponding to $w^{(j)}$. Put $F' := \delta^{-1}F$, and let $F'^{(j)}$ be the polynomial obtained by taking the j-th conjugates of the coefficients of F'.

We keep the argument together for both hyper- and superelliptic equations by using the worse bounds everywhere. Let $x, y \in B$ be a solution of (3.6), where $m, n \ge 2$ and $n \ge 3$ if $m=2$. Then

$$F'^{(j)}(x^{(j)}) = (y^{(j)})^m, \quad x^{(j)}, y^{(j)} \in \mathcal{O}_{S_i}.$$

By combining Propositions 7.8 and 7.9 we obtain the generous bound

$$H_{M_i}(x^{(j)}),\ mH_{M_i}(y^{(j)}) \ \le 80n^2\big(H_{M_i}(F'^{(j)}) + g_{M_i/\Bbbk_i} + |S_i|\big).$$

For $H_{M_i}(F'^{(j)})$, $g_{M_i/\Bbbk_i}$, $|S_i|$ we have precisely the same estimates as (7.29), (7.30), (7.31). Then a similar computation as in the proof of (7.11) leads to

$$H_{M_i}(x^{(j)}),\ mH_{M_i}(y^{(j)}) \ \le \Delta_i(nd)^{\exp O(r)}. \tag{7.43}$$

Now employing Lemma 6.6 and ignoring for the moment m we get similarly as in the proof of (7.11),

$$\overline{\deg}\, x,\ \overline{\deg}\, y \le (nd)^{\exp O(r)}.$$

It remains to estimate $m\overline{\deg}\, y$. If $y \in \overline{\mathbb{Q}}$ we have $\overline{\deg}\, y = 0$. Assume that $y \notin \overline{\mathbb{Q}}$. Then $y \notin \Bbbk_i$ for at least one index i. Since $y \in B \subset \Bbbk_i(z_i, w)$ and $[\Bbbk_i(z_i,w) : \Bbbk_i(z_i)] \le D$, we have

$$H_{M_i}(y) = [M_i : \Bbbk_i(z_i,w)]H_{\Bbbk_i(z_i,w)}(y) \ge [M_i : \Bbbk_i(z_i,w)] \ge \Delta_i/D.$$

Together with (7.43) and $D \le d^r$ this implies

$$m \le (nd)^{\exp O(r)}.$$

This concludes the proof of (7.14). □

7.3 Specializations

In this section we shall consider specialization homomorphisms from the domain B to $\overline{\mathbb{Q}}$, and using these specializations together with earlier results concerning our equations in the number field case we shall finish the proof of Propositions 7.3 and 7.4.

We recall some notation. The set of places of $\mathbb{Q}$ is $\mathcal{M}_{\mathbb{Q}} = \{\infty\} \cup \{\text{primes}\}$. By $|\cdot|_\infty$ we denote the ordinary absolute value on $\mathbb{Q}$ and by $|\cdot|_p$ (p prime) the p-adic absolute value with $|p|_p = p^{-1}$. More generally, let L be an algebraic number field with set of places $\mathcal{M}_L$. Given $v \in \mathcal{M}_L$, we define the absolute value $|\cdot|_v$ in such a way that its restriction to $\mathbb{Q}$ is $|\cdot|_p$ if v lies above $p \in \mathcal{M}_{\mathbb{Q}}$. These absolute values satisfy the product formula

$$\prod_{v \in \mathcal{M}_L} |\alpha|_v^{d_v} = 1 \qquad \text{for} \quad \alpha \in L^*,$$

where $d_v := [L_v : \mathbb{Q}_p]/[L : \mathbb{Q}]$, with $p \in \mathcal{M}_{\mathbb{Q}}$ the place below v, and $\mathbb{Q}_p$, L_v the completions of $\mathbb{Q}$ at p, L at v. Note that we have $\sum_{v|p} d_v = 1$ for every $p \in \mathcal{M}_{\mathbb{Q}}$. The absolute logarithmic height of $\alpha \in L$ is defined by

$$h(\alpha) := \log \prod_{v \in \mathcal{M}_L} \max(1, |\alpha|_v^{d_v}).$$

This depends only on α and not on the choice of the number field L containing α, hence it defines a height on $\overline{\mathbb{Q}}$. For properties of the height we refer to Bombieri and Gubler [18].

Lemma 7.10. *Let $m \geq 1$ and let $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Q}}$ be distinct, and suppose that $G(X) := \prod_{j=1}^m (X - \alpha_j) \in \mathbb{Z}[X]$. Let $q, p_0, \ldots, p_{m-1}$ be integers with $\gcd(q, p_0, \ldots, p_{m-1}) = 1$ and put*

$$\beta_j := \sum_{i=0}^{m-1} \frac{p_j}{q} \alpha_j^i, \qquad j = 1, \ldots, m.$$

Then

$$\log \max(|q|, |p_0|, \ldots, |p_{m-1}|) \leq 2m^2 + (m-1)h(G) + \sum_{j=1}^m h(\beta_j).$$

Proof. This is Lemma 5.2 in Evertse and Győry [32]. □

We now consider our specializations $B \mapsto \overline{\mathbb{Q}}$ and prove some of their properties. These specializations were introduced by Győry [39] and [40] and, in a refined form, by Evertse and Győry [32].

We assume $q > 0$ and apart from that keep the notation and assumptions from Section 7.1. In particular, $K_0 := \mathbb{Q}(z_1, \ldots, z_q)$, $K := \mathbb{Q}(z_1, \ldots, z_q, w)$, $A_0 := \mathbb{Z}[z_1, \ldots, z_q]$. Further,

$B := \mathbb{Z}[z_1, \dots, z_q, w, g^{-1}]$ where g is a non-zero element of A_0 with the properties specified in Proposition 7.2, and w is integral over A_0 and has minimal polynomial

$$\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X]$$

over K_0 as in Proposition 6.1 (i). In the case $D = 1$ we take $w = 1$, $\mathcal{F}(X) = X - 1$.

Let $\mathbf{u} = (u_1, \dots, u_q) \in \mathbb{Z}^q$. Then the substitution $z_1 \to u_1, \dots, z_q \to u_q$ defines a ring homomorphism (specialization) from K_0 to $\mathbb{Q}$

$$\varphi_{\mathbf{u}} : \alpha \mapsto \alpha(\mathbf{u}) : \left\{ \alpha = \frac{g_1}{g_2} : g_1, g_2 \in A_0, g_2(\mathbf{u}) \neq 0 \right\} \to \mathbb{Q}.$$

To extend this to a ring homomorphism from B to $\overline{\mathbb{Q}}$ we have to impose some restrictions on $\mathbf{u}$. Let $\Delta_{\mathcal{F}}$ be the discriminant of $\mathcal{F}$ (with $\Delta_{\mathcal{F}} = 1$ if $D = 1$), and let

$$\mathcal{H} := \Delta_{\mathcal{F}} \cdot \mathcal{F}_D \cdot g. \tag{7.44}$$

Put

$$\begin{cases} d_0^* := \max(\deg \mathcal{F}_1, \dots, \deg \mathcal{F}_D), & d_1^* := \max(d_0^*, \deg g) \\ h_0^* := \max(h(\mathcal{F}_1), \dots, h(\mathcal{F}_D)), & h_1^* := \max(h_0^*, h(g)). \end{cases} \tag{7.45}$$

Clearly $\mathcal{H} \in A_0$ and since $\Delta_{\mathcal{F}}$ is a homogeneous polynomial in $\mathcal{F}_1, \dots, \mathcal{F}_D$ of degree $2D-2$, we have

$$\deg \mathcal{H} \leq (2D-1)d_0^* + d_1^*. \tag{7.46}$$

Further, by Proposition 6.1 (i), Proposition 7.2 and (3.4) we also have

$$\begin{cases} d_0^* \leq (2d)^{\exp O(r)}, & h_0^* \leq (2d)^{\exp O(r)}(h+1), \\ d_1^* \leq (nd)^{\exp O(r)}, & h_1^* \leq (nd)^{\exp O(r)}(h+1) \end{cases} \tag{7.47}$$

Next assume that

$$\mathcal{H}(\mathbf{u}) \neq 0. \tag{7.48}$$

Then we have $g(\mathbf{u}) \neq 0$, $\Delta_F(\mathbf{u}) \neq 0$, hence the polynomial

$$\mathcal{F}_{\mathbf{u}} := X^D + \mathcal{F}_1(\mathbf{u}) X^{D-1} + \cdots + \mathcal{F}_D(\mathbf{u})$$

has D distinct zeros which are all different from 0, say $w^{(1)}(\mathbf{u}), \dots, w^{(D)}(\mathbf{u})$. Consequently, for $j = 1, \dots, D$ the assignment

$$z_1 \mapsto u_1, \dots, z_q \mapsto u_q, w \mapsto w^{(j)}(\mathbf{u})$$

defines a ring homomorphism $\varphi_{\mathbf{u},j}$ from B to $\overline{\mathbb{Q}}$; if $D = 1$ it is just $\varphi_{\mathbf{u}}$. The image of $\alpha \in B$ under $\varphi_{\mathbf{u},j}$ is denoted by $\alpha^{(j)}(\mathbf{u})$. It is important to note that if α is a unit in B, then its image by a specialization cannot be 0. Thus by Proposition 7.2, $\delta(\mathbf{u}) \neq 0$ and $D_F(\mathbf{u}) \neq 0$.

Recall that we may express elements of B as

$$\alpha = \sum_{i=1}^{D-1} (P_i/Q)\, w^i \qquad (7.49)$$
$$\text{where } P_0, \ldots, P_{D-1}, Q \in A_0,\ \gcd(P_0, \ldots, P_{D-1}, Q) = 1.$$

Because of $\alpha \in B$, Q must divide a power of g; hence $Q(\mathbf{u}) \neq 0$. So we have

$$\alpha^{(j)}(\mathbf{u}) = \sum_{i=1}^{D-1} (P_i(\mathbf{u})/Q(\mathbf{u})) \left(w^{(j)}(\mathbf{u})\right)^i, \qquad j = 1, \ldots, D. \qquad (7.50)$$

Clearly, $\varphi_{\mathbf{u},j}$ is the identity on $B \cap \mathbb{Q}$. Hence if $\alpha \in B \cap \overline{\mathbb{Q}}$ then $\varphi_{\mathbf{u},j}(\alpha)$ has the same minimal polynomial as α and so it is a conjugate of α.

For $\mathbf{u} = (u_1, \ldots, u_q) \in \mathbb{Z}^q$, put $|\mathbf{u}| := \max(|u_1|, \ldots, |u_q|)$. It is easy to check that for any $g \in A_0$, $\mathbf{u} \in \mathbb{Z}^q$

$$\log |g(\mathbf{u})| \leq q \log \deg g + h(g) + \deg g \log \max(1, |\mathbf{u}|). \qquad (7.51)$$

In particular, we have

$$h(\mathcal{F}_{\mathbf{u}}) \leq q \log d_0^* + h_0^* + d_0^* \log \max(1, |\mathbf{u}|) \qquad (7.52)$$

and so by Lemma 5.1 of Evertse and Győry [32]

$$\sum_{j=1}^{D} h(w^{(j)}(\mathbf{u})) \leq D + 1 + q \log d_0^* + h_0^* + d_0^* \log \max(1, |\mathbf{u}|). \qquad (7.53)$$

We define the algebraic number fields $K_{\mathbf{u},j} = \mathbb{Q}(w^{(j)}(\mathbf{u}))$ for $j = 1, \ldots, D$. We denote by Δ_L the the discriminant of an algebraic number field L. We derive an upper bound for the absolute value of the discriminant $\Delta_{K_{\mathbf{u},j}}$ of $K_{\mathbf{u},j}$.

We recall that Lemma 6.8 provides an upper bound for $|\Delta_{K_{\mathbf{u},j}}|$, Lemma 6.9 bounds the height of $\alpha^{(j)}(\mathbf{u})$ for $\mathbf{u} \in \mathbb{Z}^q$ in terms of the size of $\alpha \in B$ and some parameters of B and Lemma 6.10 shows that if we take a sufficiently large number of specializations, then there is at least one specialization among them (say corresponding to $\mathbf{u} \in \mathbb{Z}^q$), such that $\overline{h}(\alpha)$ for $\alpha \in B$ can be bounded by the heights of the images of α by the specializations $\varphi_{\mathbf{u},j}$ for $j = 1, \ldots, D$.

7.4 Bounding the height and the exponent m

We shall derive the height bounds (7.12) in Proposition 7.3 and (7.15) in Proposition 7.4, as well as the upper bound for m in Proposition 7.5 by combining the specialization techniques from the previous section with existing effective results for Diophantine equations over S-integers of a number field, namely Győry and Yu [41] for Thue equations, and the three authors [9] for hyper- and superelliptic equations and the Schinzel-Tijdeman equation.

7.4.1 Thue equations

To state the result of Győry and Yu we need some notation.

For an algebraic number field L, we denote by d_L, $\mathcal{O}_L$, $\mathcal{M}_L$, Δ_L, h_L, r_L and R_L the degree, ring of integers, set of places, discriminant, class number, unit rank and regulator of L. The absolute norm of an ideal $\mathfrak{a}$ of $\mathcal{O}_L$ is denoted by $N(\mathfrak{a})$.

Let L be an algebraic number field and let S be a finite set of places of L which contains all infinite places. Denote by s the cardinality of S. Recall that the ring of S-integers $\mathcal{O}_S$ is defined as

$$\mathcal{O}_S = \{\alpha \in L \ : \ |\alpha|_v \leq 1 \text{ for } v \in \mathcal{M}_L \setminus S\}.$$

If S consists only of the infinite places of L, we put $P := 2, Q := 2$. If S contains also finite places, we denote by $\mathfrak{p}_1, \ldots, \mathfrak{p}_t$ the prime ideals corresponding to the finite places of S, and we put

$$P := \max(N(\mathfrak{p}_1), \ldots, N(\mathfrak{p}_t)), \qquad Q := N(\mathfrak{p}_1 \ldots \mathfrak{p}_t).$$

The S-regulator associated with S is denoted by R_S. If S consists only of the infinite places of L it is just R_L, while otherwise

$$R_S = h_S R_L \prod_{i=1}^{t} \log N(\mathfrak{p}_i),$$

where h_S is a (positive) divisor of h_L. It is an easy consequence of formula (2) of Louboutin [53] that

$$h_L R_L \leq |\Delta_L|^{1/2} (\log^* |\Delta_L|)^{d_L - 1}; \tag{7.54}$$

cf. formula (59) of Győry and Yu, [41]. Further, we have

$$R_S \leq |\Delta_L|^{1/2} (\log^* |\Delta_L|)^{d_L - 1} (\log^* Q)^s; \tag{7.55}$$

see (6.1) in Evertse and Győry [32]. In view of (7.54) this is true also if $t = 0$.

7.4.1.1 Results in the number field case

Let $F(X,Y) \in L[X,Y]$ be a binary form of degree $n \geq 3$ with splitting field L and with at least three pairwise non-proportional linear factors. Further, let $\beta \in L \setminus \{0\}$ and consider the Thue equation

$$F(\xi,\eta) = \beta \qquad \text{in} \quad \xi,\eta \in \mathcal{O}_S. \tag{7.56}$$

For a polynomial G with algebraic coefficients, we denote by $h(G)$ the maximum of the logarithmic heights of its coefficients.

Proposition 7.11. *All solutions $(\xi,\eta) \in \mathcal{O}_S^2$ of equation (7.56) satisfy*

$$\begin{aligned}\max(h(\xi),h(\eta)) \leq c_1 P R_S\,(1+(\log^* R_S)/\log^* P)\times \\ \times\left(c_2 R_L + \frac{h_L}{d_L}\log Q + 2nd_L H_1 + H_2\right),\end{aligned} \tag{7.57}$$

where

$$H_1 = \max(1,h(F)), \quad H_2 = \max(1,h(\delta)),$$

$$c_1 = 250 n^6 s^{2s+3.5}\cdot 2^{7s+27}(\log 2s) d_L^{2s+4}(\log^*(2d_L))^3$$

and

$$c_2 = \begin{cases} 0 & \text{if } r_L = 0 \\ 1/d_L & \text{if } r_L = 1 \\ 29 e r_L! r_L \sqrt{r_L - 1}\log d_L & \text{if } r_L \geq 2. \end{cases}$$

Proof. This is Corollary 3 of Győry and Yu [41]. □

We shall also need the following.

Lemma 7.12. *If L is the compositum of the algebraic number fields $L_1,\ldots,L_k$ with degrees $d_{L_1},\ldots,d_{L_k}$ and discriminants $\Delta_{L_1},\ldots,\Delta_{L_k}$, then Δ_L divides $\Delta_{L_1}^{d_L/d_{L_1}}\ldots\Delta_{L_k}^{d_L/d_{L_k}}$ in $\mathbb{Z}$.*

Proof. See Stark [72]. □

Lemma 7.13. *Let L be an algebraic number field and θ a zero of a polynomial $G \in L[X]$ of degree n without multiple roots. Then*

$$|\Delta_{L(\theta)}| \leq n^{(2n-1)d_L} e^{(2n^2-2)h(G)} |\Delta_L|^{[L(\theta):L]}.$$

Proof. This is a slight modification of the second assertion of [9, Lemma 4.1]. In fact, this lemma gives the same bound but with an exponent $(2n-2)h'(G)$ on e, where for $G = \sum_{k=0}^{n} b_k X^{n-k}$ we define

$$h'(G) = \sum_{v \in \mathcal{M}_L} d_v \log \max(1, |b_0|_v, \ldots, |b_n|_v).$$

This height is easily estimated from above by $\sum_{k=0}^{n} h(b_k) \leq (n+1)h(G)$. Our lemma follows. □

7.4.1.2 Concluding the proof of Proposition 7.3

Proof of (7.12) in Proposition 7.3. We first consider the case $q > 0$. Let x, y be a solution of (7.10) in B. We keep the notation introduced in Section 7.3. Recall that $\mathcal{H} := \Delta_{\mathcal{F}} \cdot \mathcal{F}_D \cdot g$. From (7.46) and (7.47) we easily deduce

$$\deg \mathcal{H} \leq (nd)^{\exp O(r)}. \tag{7.58}$$

Choose $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$, choose $j \in \{1, \ldots, D\}$, and denote by $F_{\mathbf{u},j}$, $\delta^{(j)}(\mathbf{u})$, $x^{(j)}(\mathbf{u})$, $y^{(j)}(\mathbf{u})$, the images of F, δ, x, y under $\varphi_{\mathbf{u},j}$. Then $F_{\mathbf{u},j}$ has its coefficients in $K_{\mathbf{u},j}$. Further, let L denote the splitting field of $F_{\mathbf{u},j}$ over $K_{\mathbf{u},j}$, and S the set of places of L which consists of all infinite places and all finite places lying above the rational prime divisors of $g(\mathbf{u})$. Note that $w^{(j)}(\mathbf{u})$ is an algebraic integer and $g(\mathbf{u}) \in \mathcal{O}_S^*$. Thus $\varphi_{\mathbf{u},j}(B) \subseteq \mathcal{O}_S$ and it follows from (7.10) that

$$F_{\mathbf{u},j}\left(x^{(j)}(\mathbf{u}), y^{(j)}(\mathbf{u})\right) = \delta^{(j)}(\mathbf{u}), \qquad x^{(j)}(\mathbf{u}), y^{(j)}(\mathbf{u}) \in \mathcal{O}_S. \tag{7.59}$$

We already proved in Section 7.2 that (7.11) of Proposition 7.3 holds, i.e. we have

$$\overline{\deg}\, x, \overline{\deg}\, y \leq (nd)^{\exp O(r)}.$$

Hence we can apply Lemma 6.10 with

$$N = \max\left((nd)^{\exp O(r)}, 2Dd_0^* + 2(q+1)(d_1^*+1)\right).$$

In view of (7.47), $D \leq d^r$ and $q \leq r$ we get

$$N \leq (nd)^{\exp O(r)}. \tag{7.60}$$

By applying Lemma 6.10 with $\alpha = x$ and $\alpha = y$, and inserting $D \leq d^r$ and the upper bound $h_1^* \leq (nd)^{\exp O(r)}(h+1)$ from (7.47), it follows that there are $\mathbf{u} \in \mathbb{Z}^q$, $j \in \{1, \ldots, D\}$ with

$$|\mathbf{u}| \leq (nd)^{\exp O(r)}, \quad \mathcal{H}(\mathbf{u}) \neq 0 \tag{7.61}$$

and

$$\max(\overline{h}(x),\overline{h}(y)) \le (nd)^{\exp O(r)}\Big[(h+1)^2 + \\ +d^r(h+1)\max\big(h(x^{(j)}(\mathbf{u})),h(y^{(j)}(\mathbf{u}))\big)\Big]. \tag{7.62}$$

We proceed further with this $\mathbf{u}$, j and apply Proposition 7.11 to equation (7.59) to derive an upper bound for $h(x^{(j)}(\mathbf{u}))$ and $h(y^{(j)}(\mathbf{u}))$. To do so we have to bound from above the parameters corresponding to those which occur in Proposition 7.11.

Write $F=\sum_{k=0}^n a_k X^{n-k}Y^k$ and put

$$\overline{\deg}\,F := \max_{0\le k\le n}\overline{\deg}\,a_k,\quad \overline{h}(F) := \max_{0\le k\le n}\overline{h}(a_k).$$

Notice that by Lemma 6.3, applied to δ and the coefficients of F with the choice $d^*=d$, $h^*=h$, we have

$$\overline{\deg}\,F, \overline{\deg}\,\delta \le (2d)^{\exp O(r)}, \tag{7.63}$$

$$\overline{h}(F), \overline{h}(\delta) \le (2d)^{\exp O(r)}(h+1). \tag{7.64}$$

It follows from Lemma 6.9, $q\le r$, $D\le d^r$, (7.47), (7.63), (7.64), and lastly (7.61), that

$$\begin{aligned} h(F_{\mathbf{u},j}) &\le D^2 + q(D\log d_0^* + \log\overline{\deg}\,F) + Dh_0^* + \\ &\qquad +\overline{h}(F) + (Dd_0^* + \overline{\deg}\,F)\log\max(1,|\mathbf{u}|) \\ &\le (nd)^{\exp O(r)}(h+1). \end{aligned} \tag{7.65}$$

In a similar way, replacing F by δ, we obtain also

$$h(\delta^{(j)}(\mathbf{u})) \le (nd)^{\exp O(r)}(h+1). \tag{7.66}$$

We recall that d_L and Δ_L denote the degree and the discriminant of L over $\mathbb{Q}$. Since $[K_{\mathbf{u},j}:\mathbb{Q}]\le D$, we have $d_L\le Dn!$. Let $G(X):=F(X,1)$, and let $\theta_1,\dots,\theta_{n'}$ be the roots of G. We have $n'=n$ if $a_0\ne 0$ and $n'=n-1$ otherwise. Then $L=K_{\mathbf{u},j}(\theta_1,\dots,\theta_{n'})$. Denote by d_{L_i} the degree and by Δ_{L_i} the discriminant of the number field $L_i:=K_{\mathbf{u},j}(\theta_i)$, $i=1,\dots,n'$. Then by Lemma 7.12 we have

$$|\Delta_L| \le \prod_{i=1}^{n'} |\Delta_{L_i}|^{d_L/d_{L_i}}. \tag{7.67}$$

We estimate $|\Delta_L|$. First notice that by Lemma 6.8, inserting the estimates $q\le r$, $D\le d^r$, (7.47), (7.61),

$$\begin{aligned} |\Delta_{K_{\mathbf{u},j}}| &\le D^{2D-1}\big((d_0^*)^q e^{h_0^*}\max(1,|\mathbf{u}|^{d_0^*}|)\big)^{2D-2} \\ &\le \exp\big((nd)^{\exp O(r)}(h+1)\big). \end{aligned} \tag{7.68}$$

Further, by Lemma 7.13 and the estimates $D \le d^r$, (7.65), (7.68),

$$\begin{aligned} |\Delta_{L_i}| &\le n^{(2n-1)D} e^{(2n^2-2)h(F_{\mathbf{u},j})} |\Delta_{K_{\mathbf{u},j}}|^{[L_i:K_{\mathbf{u},j}]} \\ &\le \exp\{[L_i : K_{\mathbf{u},j}] \cdot (nd)^{\exp O(r)}(h+1)\}. \end{aligned}$$

By inserting this into (7.67), using $[L : K_{\mathbf{u},j}] \le n!$, we obtain

$$\begin{aligned} |\Delta_L| &\le \exp\left\{(nd)^{\exp O(r)}(h+1) \cdot nd_L/d_{K_{\mathbf{u},j}}\right\} \\ &\le \exp\{n!(nd)^{\exp O(r)}(h+1)\}. \end{aligned} \tag{7.69}$$

By assumption (7.45), g has degree at most d_1^* and logarithmic height at most h_1^*. Further, $g(\mathbf{u}) \neq 0$ and by $q \le r$, (7.47), (7.61),

$$|g(\mathbf{u})| \le (d_1^*)^q e^{h_1^*} \max(1, |\mathbf{u}|)^{d_1^*} \le \exp\{(nd)^{\exp O(r)}(h+1)\}. \tag{7.70}$$

The cardinality s of S is at most $d_L(1+\omega)$, where ω denotes the number of distinct prime divisors of $f(\mathbf{u})$. By prime number theory,

$$s = O(d_L \log^* |g(\mathbf{u})| / \log^* \log^* |g(\mathbf{u})|). \tag{7.71}$$

From this estimate and (7.70), $D \le d^r$, $d_L \le n!d^r$, one easily deduces that for c_1 coming from Proposition 7.11 we have

$$c_1 \le \exp\{n!(nd)^{\exp O(r)}(h+1)\}. \tag{7.72}$$

Next, we estimate P, Q and R_S. By (7.70), $d_L \le n!d^r$ we have

$$P \le Q \le |g(\mathbf{u})|^{d_L} \le \exp\{n!(nd)^{\exp O(r)}(h+1)\}. \tag{7.73}$$

To estimate R_S, we use (7.55). Then, in view of (7.69) and $d_L \le n!d^r$, we have

$$|\Delta_L|^{1/2} (\log^* |\Delta_L|)^{d_L - 1} \le \exp\{n!(nd)^{\exp O(r)}(h+1)\}. \tag{7.74}$$

Further, by (7.71) and (7.73),

$$(\log Q)^s \le \exp\left\{ O\Big(d_L \frac{\log^* |g(\mathbf{u})|}{\log^* \log^* |g(\mathbf{u})|} \cdot (\log d_L + \log^* \log^* |g(\mathbf{u})|) \Big) \right\}.$$

Together with (7.70), this leads to

$$R_S \le |\Delta_L|^{1/2} (\log^* |\Delta_L|)^{d_L - 1} (\log Q)^s \le \exp\{n!(nd)^{\exp O(r)}(h+1)\}. \tag{7.75}$$

Combining (7.54) with (7.74) and with $R_L > 0.2052$ (see Friedman [35]) we get

$$\max(h_L, R_L) \leq \exp\{n!(nd)^{\exp O(r)}(h+1)\}. \tag{7.76}$$

Finally, using $r_L < d_L \leq n!d^r$, we infer that

$$c_2 \leq \exp O(d_L \log^* d_L) \leq \exp\{n!(nd)^{\exp O(r)}\}. \tag{7.77}$$

We now apply Proposition 7.11 to equation (7.59). From the estimates (7.65), (7.66), (7.72), (7.73), (7.75), (7.76), (7.77), it follows that the upper bound in Proposition 7.11 is a sum and product of terms, which are all bounded above by $\exp\{n!(nd)^{\exp O(r)}(h+1)\}$. It follows that

$$h\left(x^{(j)}(\mathbf{u})\right), h\left(y^{(j)}(\mathbf{u})\right) \leq \exp\{n!(nd)^{\exp O(r)}(h+1)\}.$$

By inserting this into (7.62), we obtain the upper bound (7.12) in Proposition 7.3 for $q > 0$.

Now assume $q = 0$. In this case $K_0 = \mathbb{Q}$, $A_0 = \mathbb{Z}$ and $B = \mathbb{Z}[w, g^{-1}]$, where w is an algebraic integer with minimal polynomial $\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in \mathbb{Z}[X]$ over $\mathbb{Q}$, and g is a non-zero rational integer. In view of Propositions 6.1 (i) and 7.2 we may assume that

$$\log|g| \leq h_1^* \qquad \text{and} \qquad \log|\mathcal{F}_k| \leq h_0^* \qquad \text{for} \quad k = 1, \ldots, D,$$

where h_0^*, h_1^* satisfy (7.47). Denote by $w^{(1)}, \ldots, w^{(D)}$ the conjugates of w, and let $K_j := \mathbb{Q}(w^{(j)})$ for $j := 1, \ldots, D$. By a similar argument as in the proof of Lemma 5.5 of Evertse and Győry [32], we have $|\Delta_{K_j}| \leq D^{2D-1}e^{(2D-2)h_0^*}$, which is the estimate from Lemma 6.8 with $q = 0$ and $\max(1, |\mathbf{u}|)$ replaced by 1. For $\alpha \in K$, we denote by $\alpha^{(j)}$ the conjugate of α corresponding to $w^{(j)}$.

Instead of Lemma 6.10 we use Lemma 7.10, applied with $G = \mathcal{F}$, $m = D$ and $\beta^{(j)} = x^{(j)}$, resp. $y^{(j)}$. Inserting (7.47), this leads to an estimate

$$\max(\overline{h}(x), \overline{h}(y)) \leq (nd)^{\exp O(r)} \max_{1 \leq j \leq D} \max\left(h(x^{(j)}), h(y^{(j)})\right). \tag{7.78}$$

We proceed further with the j for which the maximum is assumed.

Now we can follow the argument for the case $q > 0$, except that in all estimates we have to take $q = 0$, and replace $\max(1, |\mathbf{u}|)$ by 1, $K_{\mathbf{u},j}$ by K_j, $f(\mathbf{u})$ by f, $F_{\mathbf{u},j}$ by $F^{(j)}$, where $F^{(j)}$ is the binary form obtained by taking the j-th conjugates of the coefficients of F, and $g(\mathbf{u})$ by g. This leads to an estimate

$$h((x^{(j)})), h((y^{(j)}) \leq \exp\{n!(nd)^{\exp O(r)}(h+1)\},$$

and combined with (7.78) this gives again (7.12). This completes the proof of Proposition 7.3. □

7.4.2 Hyper- and superelliptic equations

7.4.2.1 Results in the number field case.

Let L be a number field, and denote as usual by d_L, Δ_L, $\mathcal{O}_L$, $\mathcal{M}_L$ its degree, discriminant, class number, regulator, ring of integers, and set of places. Further, let S be a finite set of places of L containing all infinite places. If S consists only of the infinite places of L, put $P := 2, Q := 2$. Otherwise, denote by $\mathfrak{p}_1, \ldots, \mathfrak{p}_t$ the prime ideals corresponding to the finite places of S, and put

$$P := \max(N(\mathfrak{p}_1), \ldots, N(\mathfrak{p}_t)), \qquad Q := N(\mathfrak{p}_1 \ldots \mathfrak{p}_t).$$

Let

$$F(X) = a_0 X^n + a_1 X^{n-1} + \cdots + a_n \in \mathcal{O}_S[X] \tag{7.79}$$

be a polynomial of degree $n \geq 2$ and of non-zero discriminant, $\delta \in \mathcal{O}_S \setminus \{0\}$, and m a positive integer. Put

$$\hat{h} := \sum_{v \in \mathcal{M}_L} d_v \log \max(1, |\delta|_v, |a_0|_v, \ldots, |a_n|_v),$$

where $d_v := [L_v : \mathbb{Q}_p]/[L : \mathbb{Q}]$, with $p \in \mathcal{M}_{\mathbb{Q}}$ the place below v.

Proposition 7.14. *Assume $n \geq 2$, $m \geq 3$. If $x, y \in \mathcal{O}_S$ is a solution to the equation*

$$F(x) = \delta y^m, \qquad x, y \in \mathcal{O}_S, \tag{7.80}$$

then

$$h(x), h(y) \leq c_3^{m^3} |\Delta_L|^{2m^2n^2} Q^{3m^2n^2} e^{8m^2n^3 d_L \widehat{h}},$$

where $c_3 := (6ns)^{14n^3s}$.

Proof. This is Theorem 2.1 in [9]. □

Proposition 7.15. *Let $n \geq 3$. If $x, y \in \mathcal{O}_S$ is a solution to*

$$F(x) = \delta y^2, \qquad x, y \in \mathcal{O}_S, \tag{7.81}$$

then

$$h(x), h(y) \leq c_4 |\Delta_L|^{8n^3} Q^{20n^3} e^{50n^4 d_L \widehat{h}},$$

where $c_4 := (4ns)^{212n^4s}$.

Proof. This is Theorem 2.2 in [9]. □

Proposition 7.16. *Let $n \geq 2$. If x, y, m is a solution to*

$$F(x) = \delta y^m, \qquad x, y \in \mathcal{O}_S,\ m \in \mathbb{Z}_{\geq 2},$$

such that $y \neq 0$ and y is not a root of unity, then

$$m \leq c_5 |\Delta_L|^{6n} P^{n^2} e^{11nd_L \widehat{h}},$$

where $c_5 := (10n^2 s)^{40ns}$.

Proof. This is Theorem 2.3 in [9]. □

7.4.2.2 Concluding the proofs of Propositions 7.4 and 7.5

Proof of (7.15) in Proposition 7.4. The computations will be similar to those in the proof of (7.12) in Proposition 7.3 but with some simplifications.

First we suppose $q > 0$. Take a solution x, y of (7.13) in B. We use again the polynomial $\mathcal{H} := \Delta_{\mathcal{F}} \cdot \mathcal{F}_D \cdot g$ from Section 7.3. Take again $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$, choose $j \in \{1, \ldots, D\}$, and denote by $F_{\mathbf{u},j}$, $\delta^{(j)}(\mathbf{u})$, $x^{(j)}(\mathbf{u})$, $y^{(j)}(\mathbf{u})$, the images of F, δ, x, y under the specialization $\varphi_{\mathbf{u},j}$. In contrast to our argument for Thue equations, we do not have to deal with the splitting field of F now. So we take for S the set of places of $K_{\mathbf{u},j}$, consisting of all infinite places, and all finite places lying above the rational prime divisors of $g(\mathbf{u})$. Then $\varphi_{\mathbf{u},j}(B) \subseteq \mathcal{O}_S$, and

$$F_{\mathbf{u},j}(x^{(j)}(\mathbf{u})) = \delta^{(j)}(\mathbf{u}) y^{(j)}(\mathbf{u})^m, \quad \mathbf{x}^{(j)}(\mathbf{u}),\ y^{(j)}(\mathbf{u}) \in \mathcal{O}_S. \tag{7.82}$$

Note that by the choice of $\mathcal{H}$ and $\mathcal{H}(\mathbf{u}) \neq 0$ we have $\delta_j(\mathbf{u}) \neq 0$ and $F_{\mathbf{u},j}$ has non-zero discriminant. So $F_{\mathbf{u},j}$ has the same number of zeros and the same degree as F, that is, the degree of $F_{\mathbf{u},j}$ is $n \geq 2$ if $m \geq 3$ and $n \geq 3$ if $m = 2$. Hence Propositions 7.14 and 7.15 are applicable.

By precisely the same argument as in the case for Thue equations, there are $\mathbf{u} \in \mathbb{Z}^q$ and $j \in \{1, \ldots, D\}$ with (7.61) and (7.62). We proceed further with this $\mathbf{u}$, j.

We estimate the parameters corresponding to those in the bounds from Propositions 7.14, 7.15. First, we get precisely the same estimates as in (7.65) and (7.66). These imply

$$\widehat{h} \leq (n+1) h(F_{\mathbf{u},j}) + h(\delta^{(j)}(\mathbf{u})) \leq (nd)^{\exp O(r)} (h+1). \tag{7.83}$$

Further we have, similarly to (7.68),

$$|\Delta_{K_{\mathbf{u},j}}| \leq \exp\{(nd)^{\exp O(r)} (h+1)\}. \tag{7.84}$$

Next, similar to (7.70),

$$|g(\mathbf{u})| \le \exp\{(nd)^{\exp O(r)}(h+1)\}. \tag{7.85}$$

The set S now consists of places of $K_{\mathbf{u},j}$ instead of the splitting field of $F_{\mathbf{u},j}$ over K. So since $[K_{\mathbf{u},j} : \mathbb{Q}] \le D$ we now have $s \le D(1+\omega)$, where ω is the number of distinct prime divisors of $g(\mathbf{u})$. This gives, instead of (7.71),

$$s = O\left(D \log^* |g(\mathbf{u})| / \log^* \log^* |g(\mathbf{u})|\right). \tag{7.86}$$

By inserting (7.85), and $D \le d^r$, we obtain for the quantities c_3, c_4 in Propositions 7.14 and 7.15 the upper bounds

$$c_3,\ c_4\ \le \exp\{(nd)^{\exp O(r)}(h+1)\}. \tag{7.87}$$

Lastly, we have instead of (7.73),

$$P \le Q \le |g(\mathbf{u})|^D \le \exp\{(nd)^{\exp O(r)}(h+1)\}, \tag{7.88}$$

where we have used (7.85) and $D \le d^r$.

We now apply Propositions 7.14 and 7.15 to (7.82). Note that we have to take $L = K_{\mathbf{u},j}$; so $d_L \le D \le d^r$. By inserting this and (7.83), (7.84), (7.87), (7.88) into the upper bounds from these Propositions, we obtain

$$h(x^{(j)}(\mathbf{u})),\ h(y^{(j)}(\mathbf{u}))\ \le \exp\{m^3(nd)^{\exp O(r)}(h+1)\}. \tag{7.89}$$

By inserting this into (7.62), we obtain (7.15) in the case $q > 0$.

Now let $q = 0$. For $\alpha \in K$, write $\alpha^{(j)}$ for the conjugate of α corresponding to $w^{(j)}$, and let $F^{(j)}$ be the polynomial obtained by taking the j-th conjugates of the coefficients of F. We simply have to follow the above arguments, replacing everywhere q by 0, $\max(1, |\mathbf{u}|)$ by 1, $K_{\mathbf{u},j}$ by $K^{(j)} = \mathbb{Q}(w^{(j)})$, $F_{\mathbf{u},j}$ by $F^{(j)}$, $x^{(j)}(\mathbf{u})$, $y^{(j)}(\mathbf{u})$ by $x^{(j)}$, $y^{(j)}$, and $g(\mathbf{u})$ by $g \in \mathbb{Z}$. Instead of (7.62) we have to use (7.78). Thus, we obtain the same estimate as (7.89), but with $x^{(j)}$, $y^{(j)}$ instead of $x_j(\mathbf{u})$, $y_j(\mathbf{u})$. Via (7.78) we obtain (7.15) in the case $q = 0$. This completes our proof of Proposition 7.4. □

Proof of Proposition 7.5. Assume for the moment $q > 0$. Let $x \in B$, $y \in B \cap \overline{\mathbb{Q}}$, $m \in \mathbb{Z}_{\ge 2}$ be a solution of (7.13), such that $y \ne 0$ and y is not a root of unity. Choose again $\mathbf{u}$, j with (7.61), (7.62). Note that $y^{(j)}(\mathbf{u})$ is a conjugate of y since $y \in \overline{\mathbb{Q}}$; hence it is not 0 or a root of unity.

We apply Proposition 7.16 to (7.82). By (7.85), (7.86), we have for the constant c_5 in Proposition 7.16, that

$$c_5 \le \exp\{(nd)^{\exp O(r)}(h+1)\}.$$

Further, we have the upper bounds (7.83) for $\widehat{h}$, (7.84) for $|\Delta_{K_{\mathbf{u},j}}|$, and (7.88) for P. By inserting these estimates into the upper bound for m from Proposition 7.16, we obtain $m \leq \exp\{(nd)^{\exp O(r)}(h+1)\}$. In the case $q = 0$, we obtain the same estimate, by making the same modifications as in the proof of Proposition 7.4. This finishes our proof of Proposition 7.5. □

Chapter 8

Proof of the results from Section 3.3

8.1 Preparation for the proof of Theorem 3.5

8.1.1 Analyzing the condition (3.13) posed on F

Let $A, K, \overline{K}$ be as in Section 3.3 and let $F(X,Y) \in A[X,Y]$ be a bivariate polynomial given by

$$F(X,Y) = \sum_{(i,j)\in I} a_{ij} X^i Y^j,$$

where $I \subset \mathbb{Z}^2_{\geq 0}$ is a finite set, and $0 \neq a_{ij} \in A$ are fixed for $(i,j) \in I$. Denote by N the total degree of F and by $n(F)$ the number of non-zero coefficients of F.

A partition of the set I is just a tuple $\mathcal{P} = (I_1, \ldots, I_k)$ of subsets of I with the properties $I_1 \cup I_2 \cup \cdots \cup I_k = I$, $I_i \cap I_j = \emptyset$ for $i \neq j$, and $I_l \neq \emptyset$ for $l = 1, \ldots, k$.

For any partition $\mathcal{P} = (I_1, \ldots, I_k)$ of I with $|I_l| \geq 2$ for $l = 1, \ldots, k$ we define the $\mathbb{Z}$-module

$$\Lambda(F, \mathcal{P}) := \langle \{(i_1, j_1) - (i_2, j_2) \mid (i_1, j_1), (i_2, j_2) \in I_l \text{ for some } l = 1, \ldots, k\} \rangle$$

i.e. the $\mathbb{Z}$-module defined by all differences of pairs of exponents (i,j) belonging to the same set in the partition $\mathcal{P}$. Let $r(F, \mathcal{P})$ denote the rank of the $\mathbb{Z}$-module $\Lambda(F, \mathcal{P})$.

In the sequel, for any solution (x,y) of the equation

$$F(x,y) = 0 \qquad \text{in} \quad x, y \in A^* \tag{8.1}$$

we say that a partition $\mathcal{P} = (I_1, \ldots, I_k)$ of I corresponds to F and (x,y) if

i. x, y is a solution of the following system

$$\sum_{(i,j)\in I_l} a_{ij} x^i y^j = 0 \qquad \text{for} \quad l = 1, \ldots, k, \tag{8.2}$$

ii. and $\sum_{(i,j)\in I_0} a_{ij}x^iy^j \neq 0$ for any proper subset I_0 of any of the sets I_l for $l = 1, \ldots, k$.

In this case we shall also say that (x, y) is associated with the partition $\mathcal{P}$. We mention that $a_{ij} \neq 0$, $x, y \in A^*$ and (8.2) imply $|I_l| \geq 2$ for $l = 1, \ldots, k$.

In the case when for a given partition $\mathcal{P}$ the rank of $\Lambda := \Lambda(F, \mathcal{P})$ is 1 then we associate a system of polynomials to $\mathcal{P}$ as follows: if $r(F, \mathcal{P}) = 1$ then there exists a pair $(m, n) \in \mathbb{Z}^2$ with $\gcd(m, n) = 1$ such that for any two elements $(i, j), (i', j') \in I_l$ for $l = 1, \ldots, k$ we have $(i, j) - (i', j') = t \cdot (m, n)$ with $t \in \mathbb{Z}$, $|t| \leq N$. Fixing an element $(i_l, j_l) \in I_l$ for $l = 1, \ldots, k$ we get that every $(i, j) \in I_l$ can be written as $(i, j) = (i_l, j_l) + t_{ij}(m, n)$, for $l = 1, \ldots, k$, with some $t_{ij} \in \mathbb{Z}$, $|t_{ij}| \leq N$. Thus the system (8.2) is equivalent to the system

$$X^{i_l}Y^{j_l} \sum_{(i,j)\in I_l} a_{ij}(X^mY^n)^{t_{ij}} = 0 \qquad \text{for} \quad l = 1, \ldots, k.$$

By multiplying these equations by suitable powers of X^mY^n we see that it is equivalent to a system

$$g_l(X^mY^n) = 0 \qquad \text{for} \quad l = 1, \ldots, k, \tag{8.3}$$

where $g_l \in A[X]$, $g_l(0) \neq 0$ for $l = 1, \ldots, k$ and

$$g_l(X) := \sum_{(i,j)\in I_l} a_{ij}X^{s_{ij}}, \tag{8.4}$$

where $0 \leq s_{ij} \leq 2N$. We shall call $(g_1, \ldots, g_k)$ the polynomial system corresponding to the partition $\mathcal{P}$.

Now the fact that (8.1) has a solution associated with $\mathcal{P}$ is equivalent to the system (8.3) having a solution $x, y \in A^*$ which can happen only if the polynomials $g_k(X)$ have a common root $\alpha \in A^*$, i.e. $X - \alpha$ divides g_l for all $l = 1, \ldots, k$, which contradicts the assumption (3.13). Now we are ready to state two Propositions:

Proposition 8.1. *Let $F(X, Y) \in A[X, Y]$ be a polynomial. Then F satisfies condition (3.13) if and only if for any partition $\mathcal{P} = (I_1, \ldots, I_k)$ of I we have one of the following:*

i. $r(F, \mathcal{P}) = 2$, or

ii. $r(F, \mathcal{P}) = 1$, and the polynomial system $(g_1, \ldots, g_k) \in A[X]^k$ corresponding to $\mathcal{P}$ has the property

$$\gcd(g_1, \ldots, g_k) = 1 \qquad \text{in} \quad K[X].$$

Proof. First suppose that (3.13) holds. Let $\mathcal{P}$ be any partition with $r(F, \mathcal{P}) = 1$ and assume $\gcd(g_1, \ldots, g_k) \neq 1$ over K. Thus there exists $\alpha \in \overline{K}$ with $g_i(\alpha) = 0$ for $i = 1, \ldots, k$, and so $X^mY^n - \alpha$ or $X^m - \alpha Y^n$ divides F for some $m, n \in \mathbb{Z}_{\geq 0}$, which contradicts (3.13).

Conversely, we show that if F has a factor of the form $X^mY^n - \alpha$ or $X^m - \alpha Y^n$ with $m, n \in \mathbb{Z}_{\geq 0}$ and $\alpha \in \overline{K}$ then there exists a partition $\mathcal{P}$ of I such that $r(F, \mathcal{P}) = 1$ and $\gcd(g_1, \ldots, g_k) \neq 1$ over K. To simplify the proof we consider F as a Laurent polynomial. Then an equivalent formulation of our assumption is that F has a non-constant divisor of the form $X^mY^n - \alpha$ with $m, n \in \mathbb{Z}$ and $\alpha \in \overline{K}$. Clearly, we may suppose $(m, n) = 1$, thus there exist $m', n' \in \mathbb{Z}$ with $mn' - nm' = 1$. Put $U = X^mY^n$ and $V = X^{m'}Y^{n'}$, and define the Laurent polynomial F' by $F'(U, V) = F(X, Y)$. Now F' is divisible by $U - \alpha$, thus we have $F'(\alpha, V) \equiv 0$. If we write

$$F'(U, V) = \sum_{i=0}^{k} V^i g_i(U)$$

then by $F'(\alpha, V) \equiv 0$ we must have $g_i(\alpha) = 0$ for all $i = 1, \ldots, k$, and thus $\gcd(g_1, \ldots, g_k) \neq 1$ in $K[U]$. Writing F in the form

$$F(X, Y) = \sum_{i=0}^{k} X^{im'} Y^{in'} g_i(X^mY^n)$$

induces a partition $\mathcal{P} = (I_1, \ldots, I_k)$, with $r(F, \mathcal{P}) = 1$ and $\gcd(g_1, \ldots, g_k) \neq 1$ over K. This concludes the proof of the proposition. □

Proposition 8.2. *Let $F(X, Y)$ be a polynomial satisfying (3.13) and fix a solution (x, y) of (8.1). Let $\mathcal{P} = (I_1, \ldots, I_k)$ be a partition of I corresponding to F and (x, y) and let $\Lambda := \Lambda(F, \mathcal{P})$ be the $\mathbb{Z}$-module corresponding to the solution (x, y) and the partition $\mathcal{P}$. Then we have*

$$r(F, \mathcal{P}) = 2.$$

Proof. This is a direct consequence of Proposition 8.1, since for a solution (x, y) and a partition $\mathcal{P}$ with $r(F, \mathcal{P}) = 1$ associated with it, the corresponding polynomial system $g_1, \ldots, g_k$ has a gcd of degree > 0, which contradicts (ii) of Proposition 8.1. That is only $r(F, \mathcal{P}) = 2$ is possible. □

The above two propositions mean in fact, that for a polynomial fulfilling condition (3.13) there might exist partitions of I of rank 1, but these are never partitions corresponding to a solution.

8.1.2 Effective estimates for the gcd of polynomials

For a polynomial $P \in \mathbb{C}[X]$ let $||P||_1$ denote the sum of the absolute values of the coefficients of P.

Proposition 8.3. *Let A be a finitely generated domain as in Section 3.3 and K its quotient field. Let $k, \rho \in \mathbb{N}$ be with $2^{k-1} \leq \rho \leq 2^k$ and let*

$$g_i(X) := \sum_{j=0}^{\delta} x_{ij} X^j \in A[X] \qquad \text{for} \quad i = 1, \ldots, \rho$$

be non-zero polynomials such that $\gcd(g_1, \ldots, g_\rho)$ in $K[X]$ has degree δ_0. Let $\mathbf{x} := (x_{ij} : i = 1, \ldots, \rho, j = 0, \ldots, \delta)$ be the vector consisting of the coefficients of the polynomials $g_1, \ldots, g_\rho$.

Then there exist polynomials $P_0, \ldots, P_{\delta_0}$ with integer coefficients, in $\rho(\delta+1)$ variables with the following properties:

i. $\deg P_i \leq (2\delta)^k$, and $||P_i||_1 \leq (2\delta)^{2\delta+(2\delta)^2+\cdots+(2\delta)^k}$ for $i = 1, \ldots, \delta_0$;

ii. There are polynomials $u_1, \ldots, u_\rho \in A[X]$ such that

$$u_1 g_1 + \cdots + u_\rho g_\rho = \sum_{j=0}^{\delta_0} P_j(\mathbf{x}) X^j,$$

where not all $P_j(\mathbf{x})$ are 0.

For the proof of Proposition 8.3 we need the following:

Lemma 8.4. *Let A be a finitely generated domain as in Section 3.3 and K its quotient field. Let $g_1, g_2 \in A[X]$ be non-zero polynomials with $\deg g_1 = n_1$, $\deg g_2 = n_2$, and such that $\gcd(g_1, g_2)$ in $K[X]$ has degree δ_0. Then there exist polynomials $u_1, u_2, g \in A[X]$ with*

$$u_1 g_1 + u_2 g_2 = g, \tag{8.5}$$

with $\deg u_1 \leq n_2 - \delta_0 - 1$, $\deg u_2 \leq n_1 - \delta_0 - 1$, $\deg g = \delta_0$, and such that the coefficients of g are determinants of order $n_1 + n_2 - 2\delta_0$ of which $n_2 - \delta_0$ columns consist of coefficients of g_1 and $n_1 - \delta_0$ columns consist of coefficients of g_2. Further, in this case we have automatically $g = \gcd(g_1, g_2)$ in $K[X]$.

Proof. By properties of the gcd of polynomials over a field there exist $g = \gcd(g_1, g_2) \in K[X]$ and $u_1, u_2 \in K[X]$ with (8.5), and reducing u_1 modulo g_2/g, and u_2 modulo g_1/g it is clear that we may choose u_1, u_2 such that $\deg u_1 \leq n_2 - \delta_0 - 1$ and $\deg u_2 \leq n_1 - \delta_0 - 1$. Further the triple (u_1, u_2, g) is unique up to a common constant factor from K. Multiplying

the identity by a common multiple of all the denominators of the coefficients of g, u_1, u_2 we can guarantee also $g, u_1, u_2 \in A[X]$. Write

$$u_1 := \sum_{i=0}^{n_2-\delta_0-1} x_i X^i, \qquad u_2 := \sum_{i=0}^{n_1-\delta_0-1} y_i X^i \qquad g = \sum_{i=0}^{\delta_0} z_i X^i.$$

Then by equating coefficients, the polynomial identity

$$u_1 g_1 + u_2 g_2 - g = 0$$

is equivalent to a system of linear equations

$$\begin{pmatrix} -I & F_{11} & F_{12} \\ \mathbf{0} & F_{21} & F_{22} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{z} \\ \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \underline{0}.$$

in the variables x_i, y_i, z_i, consisting of $n_1 + n_2 - \delta_0$ linearly independent equations. In this system the block $-I$ is the negative of a unit matrix of order $\delta_0 + 1$, F_{11} and F_{21} are blocks (of $n_2 - \delta_0$ columns) consisting of coefficients of g_1 and F_{12} and F_{22} blocks (of $n_1 - \delta_0$ columns) consisting of coefficients of g_2.

The solution subspace of this system of equations is one-dimensional, and we have one more unknown than the number of equations. Hence the equations in the system are linearly independent. Further, this system of equations has the non-zero solution $(\Delta_1, -\Delta_2, \ldots, \pm\Delta_{n_1+n_2-\delta_0+1})^{\mathrm{T}}$, where Δ_i denotes the determinant of the matrix obtained from the matrix of our system by removing the ith column. So we may take

$$g(X) = \Delta_1 - \Delta_2 X + \Delta_3 X^2 + \cdots \pm \Delta_{\delta_0+1} X^{\delta_0}.$$

This concludes the proof of our lemma. □

Proof of Proposition 8.3. We may assume without loss of generality that $\rho = 2^k$, otherwise we copy some of the polynomials $g_1, \ldots, g_\rho$ to have 2^k polynomials.

Now we use induction on k. For $k = 1$ the statement is true by Lemma 8.4. So we assume that the statement of our proposition is true for $k-1$ and we prove it for k. Suppose that

$$\deg \gcd(g_1, \ldots, g_{2^{k-1}}) = d_1, \qquad \deg \gcd(g_{2^{k-1}+1}, \ldots, g_{2^k}) = d_2 \quad \text{in } K[X].$$

Then by the inductive assumption there are polynomials $v_1, \ldots, v_{2^{k-1}} \in A[X]$ with

$$\sum_{i=1}^{2^{k-1}} v_i g_i = \sum_{j=0}^{d_1} Q_{1j}(\mathbf{x}_1) X^j,$$

where not all Q_{1j} are zero and where $\mathbf{x}_1$ is the vector consisting of all coefficients of the polynomials $g_1, \ldots, g_{2^{k-1}}$, and there also exist polynomials $v_{2^{k-1}+1}, \ldots, v_{2^k} \in A[X]$ with

$$\sum_{i=2^{k-1}+1}^{2^k} v_i g_i = \sum_{j=0}^{d_2} Q_{2j}(\mathbf{x}_2) X^j,$$

where not all Q_{2j} are zero and where $\mathbf{x}_2$ is the vector consisting of all coefficients of the polynomials $g_{2^{k-1}+1}, \ldots, g_{2^k}$. Further, by the induction hypothesis we may assume

$$\deg Q_{ij} \le (2\delta)^{k-1}, \qquad ||Q_{ij}||_1 \le (2\delta)^{2\delta+\cdots+(2\delta)^{k-1}} := c(\delta)$$

for $i = 1, 2$ and $j = 0, \ldots, d_i$. By Lemma 8.4 there are $w_1, w_2 \in A[X]$ such that

$$w_1 \sum_{j=0}^{d_1} Q_{1j}(\mathbf{x}_1) X^j + w_2 \sum_{j=0}^{d_2} Q_{2j}(\mathbf{x}_2) X^j = \sum_{j=0}^{\delta_0} P_j(\mathbf{x}) X^j,$$

with $P_{\delta_0} \neq 0$, and where P_j is a determinant of order $d_1 + d_2 - 2\delta_0$ of which $d_2 - \delta_0$ columns consist of polynomials Q_{1j} $(j = 1, \ldots, d_1)$ and $d_1 - \delta_0$ columns of polynomials Q_{2j} $(j = 1, \ldots, d_2)$. This implies

$$\deg P_j(\mathbf{x}) \le (d_2 - \delta_0)(2\delta)^{k-1} + (d_1 - \delta_0)(2\delta)^{k-1} \le \delta(2\delta)^{k-1} + \delta(2\delta)^{k-1} \le (2\delta)^k,$$

and

$$\begin{aligned} ||P_j||_1 &\le \{(d_1 + d_2 - 2\delta_0) \cdot c(\delta)\}^{d_2-\delta_0} \cdot \{(d_1 + d_2 - 2\delta_0) \cdot c(\delta)\}^{d_1-\delta_0} \\ &\le \left\{(2\delta)^{\delta} \cdot (2\delta)^{\delta\cdot(2\delta+\cdots+(2\delta)^{k-1})}\right\}^2 \le (2\delta)^{2\delta+\cdots+(2\delta)^k}. \end{aligned}$$

This concludes the proof of Proposition 8.3. □

Corollary 8.5. *Let A be a finitely generated domain as in Section 3.3 and K its quotient field. Let $k, \rho \in \mathbb{N}$ be with $2^{k-1} \le \rho \le 2^k$ and define the polynomials*

$$g_i(X) := \sum_{j=0}^{\delta} x_{ij} X^j \in A[X] \qquad \textit{for} \quad i = 1, \ldots, \rho.$$

Further, suppose that the coefficients $x_{ij} \in A$ have representatives $\tilde{x}_{ij}$ with

$$\deg \tilde{x}_{ij} \le d, \qquad h(\tilde{x}_{ij}) \le h,$$

where $d > 1$ and $h > 1$ are given real numbers. Suppose that

$$\gcd(g_1, \ldots, g_\rho) = 1 \quad \textit{in } K[X].$$

Then there exist polynomials $u_1, \ldots, u_\rho \in A[X]$ *such that*

$$u_1 g_1 + \cdots + u_\rho g_\rho = R,$$

where $R \in A$, $R \neq 0$, *and* R *has a representative* $\tilde{R}$ *with*

$$\deg \tilde{R} \leq d(2\delta)^k, \qquad h(\tilde{R}) \leq (2\delta)^{k+2}(d+1)rh.$$

Proof. Put $\mathbf{x} := (x_{ij} \ : \ i = 1, \ldots, \rho, j = 0, \ldots, \delta)$ be the vector consisting of the coefficients of the polynomials $g_1, \ldots, g_\rho$. By Proposition 8.3 there exist polynomials $u_1, \ldots, u_\rho \in A[X]$ such that

$$u_1 g_1 + \cdots + u_\rho g_\rho = P_0(\mathbf{x}),$$

where $P_0(\mathbf{X})$ is a polynomial in $\rho(\delta+1)$ variables with integer coefficients and with

$$\deg P_0 \leq (2\delta)^k, \qquad ||P_0||_1 \leq (2\delta)^{2\delta+(2\delta)^2+\cdots+(2\delta)^k}.$$

This together with $\deg \tilde{x}_{ij} \leq d$ proves

$$\deg \tilde{R} \leq d(2\delta)^k.$$

Clearly by the assumptions of the corollary we have

$$||\tilde{x}_{ij}||_1 \leq (d+1)^r h,$$

thus

$$||\tilde{R}||_1 = ||P_0||_1((d+1)^r h)^{(2\delta)^k} \leq (2\delta(d+1)^r h)^{(2\delta)^{k+1}}.$$

and finally we get

$$h(\tilde{R}) \leq \log ||\tilde{R}||_1 \leq (2\delta)^{k+2}(d+1)rh.$$

□

8.2 Extending A to a larger ring

First we shall extend our domain A to a larger domain B and prove an effective result for the set

$$\mathcal{C}' := \{(x, y) \in (B^*)^2 | F(x, y) = 0\}$$

The main advantage of this will be, that we choose the larger domain B such that it will be easier to do effective computations with elements of B than it is with elements of A.

Recall that $A = \mathbb{Z}[z_1, \dots, z_r]$ is a finitely generated domain, and let us denote by K the quotient field of A. Let $f_1, \dots, f_t$ be the generators of the ideal $\mathcal{I}$ that defines our domain A (see (3.1), (3.2)) and put

$$d_0 := \max(1, \deg f_1, \dots, \deg f_t), \qquad h_0 := \max(1, h(f_1), \dots, h(f_t)). \tag{8.6}$$

Let $q \geq 0$ denote the transcendence degree of K and suppose without loss of generality that $z_1, \dots, z_q$ is a transcendence basis of $K/\mathbb{Q}$. Put

$$K_0 := \mathbb{Q}(z_1, \dots, z_q), \qquad A_0 := \mathbb{Z}[z_1, \dots, z_q], \tag{8.7}$$

with the convention that in the case $q = 0$ we put $K_0 = \mathbb{Q}$ and $A_0 = \mathbb{Z}$. For elements $0 \neq f \in A_0$ we will use the notation $\deg f$ and $h(f)$ for the total degree and logarithmic height of f, respectively, viewed as a polynomial in the unknowns $z_1, \dots, z_q$, with the convention that in the case $q = 0$ we put $\deg f := 0$ and $h(f) := \log |f|$.

The field K is clearly a finite algebraic extension of K_0, so we have $K = K_0(w)$ for some $w \in K$. We shall see that w may be chosen in such a way that it is integral over A_0, the degree of its minimal polynomial, and the degree and height of the coefficients of its minimal polynomial are bounded. Further, there exists an element $f \in A_0$, such that $A \subset A_0[w, f^{-1}] := B$, some "important" elements are units in B, and the degree and height of f is also bounded. This is described more precisely in the following proposition. Recall that a_{ij} denote the coefficients of F and N is the total degree of F in Theorem 3.5. Let us use the notation $\log_2^* x := \max(1, \log_2 x)$.

Proposition 8.6. *(i) There exists an element $w \in A$ which is integral over A_0 such that $K = K_0(w)$ and having minimal polynomial*

$$\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X]$$

over K_0 of degree $D \leq d_0^{r-q}$, such that

$$\deg \mathcal{F}_k \leq (2d_0)^{\exp O(r)}, \qquad h(\mathcal{F}_k) \leq (2d_0)^{\exp O(r)}(h_0 + 1) \tag{8.8}$$

for $k = 1, \dots, D$.

(ii) Let $R \in A$ and suppose that R has a representative $\tilde{R}$ with

$$\deg \tilde{R} \leq d(4N)^{\log_2^* N}, \qquad h(\tilde{R}) \leq (4N)^{\log_2^* N + 2}(d+1)rh. \tag{8.9}$$

Then there exists a non-zero $f \in A_0$ such that

$$\begin{aligned} &A \subseteq A_0[w, f^{-1}], \\ &a_{ij} \in A_0[w, f^{-1}]^* \qquad \text{for } (i,j) \in I \\ &R \in A_0[w, f^{-1}]^* \end{aligned} \tag{8.10}$$

and

$$\begin{aligned} \deg f &\le (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}, \\ h(f) &\le (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} \cdot h. \end{aligned} \tag{8.11}$$

Remark. The element R above for the moment may be any $R \in A$ with (8.9), and it will be specified at the very end of our proof in equation (8.41).

Proof of Proposition 8.6. In the proof for convenience we shall use Proposition 6.1. However, this proposition is just a suitable reformulation and combination of Proposition 3.4, Lemma 3.2, (i), and Lemma 3.6. of Evertse and Győry [32]. In principle (i) of the present proposition is exactly (i) of Proposition 3.1 of [32].

To prove (ii) we will use (ii) of Proposition 6.1 with the choice

$$\{\alpha_1, \dots, \alpha_k\} = \{a_{ij}, \quad \text{for } (i,j) \in I\} \cup \{R\}.$$

Thus we have $k = n(F) + 1 < O(N^2)$, where $n(F)$ denotes the number of non-zero coefficients of F. Further, we may choose v_l to be 1, and u_l to be one of the polynomials $\tilde{a}_{ij}$ for $l = 1, \dots, k-1$, and we also may choose $v_k = 1$ and $u_k = \tilde{R}$ which gives the estimates

$$d^{**} = d(4N)^{\log_2^* N} \quad \text{and} \quad h^{**} = (4N)^{\log_2^* N + 2}(d+1)rh.$$

Now we use statement (ii) of Proposition 6.1 and we choose a larger constant in the $O(\cdot)$ symbol to simplify the expressions in the bounds. This concludes the proof of our Proposition 8.6. □

Next we recall the representation for the elements of the field K, which we already explained in Chapter 6. As in Proposition 8.6 we denote the degree of K over K_0 by D. Since $K = K_0(w)$ every element $\alpha \in K$ can be written uniquely in the form $\sum_{j=0}^{D-1} R_{\alpha,j} w^j$, where $R_{\alpha,j} \in K_0$. Since K_0 is the fraction field of A_0, and A_0 is a unique factorization domain (indeed, $z_1, \dots, z_q$ are algebraically independent), there exist $P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha \in A_0$ such that the above representation can be rewritten in the form

$$\alpha = Q_\alpha^{-1} \sum_{j=0}^{D-1} P_{\alpha,j} w^j \quad \text{with} \quad Q_\alpha \neq 0, \quad \gcd(P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha) = 1. \tag{8.12}$$

Further, the tuple $(P_{\alpha,0}, \dots, P_{\alpha,D-1}, Q_\alpha)$ in the representation (8.12) of α is up to sign uniquely determined.

Using this representation we introduce two new concepts which will turn out to be useful to measure elements of K. Let us define

$$\begin{cases} \overline{\deg}\,\alpha := \max(\deg P_{\alpha,0}, \dots, \deg P_{\alpha,D-1}, \deg Q_\alpha) \\ \overline{h}(\alpha) := \max(h(P_{\alpha,0}), \dots, h(P_{\alpha,D-1}), h(Q_\alpha)), \end{cases} \tag{8.13}$$

with the convention that for $q = 0$ we define $\overline{h}(\alpha) = \log \max(|P_{\alpha,0}|, \dots, |P_{\alpha,D-1}|, |Q_\alpha|)$ and $\overline{\deg}\,\alpha = 0$.

Recall that Lemmas 6.3 and 6.4 show that $\overline{\deg}\,\alpha$ and $\overline{h}(\alpha)$ may be bounded by the height and degree of representatives for α, the bound being dependent also on parameters of A, and conversely, $\alpha \in A$ has a representative whose height and degree are bounded by $\overline{\deg}\,\alpha$ and $\overline{h}(\alpha)$, the bound again being dependent also on parameters of A.

In the following proposition we shall state a generalization of our Theorem 3.5 and then we show how our Theorem 3.5 follows from that. Then the rest of the Chapter will be devoted to the proof of this more general proposition.

Proposition 8.7. *Let w and f be as in Proposition 8.6 and put*

$$B := A_0[f^{-1}, w].$$

Then for every element (x, y) of the set

$$\mathcal{C}' := \{(x, y) \in (B^*)^2 | F(x, y) = 0\}$$

we have

$$\overline{\deg}\,x,\ \overline{\deg}\,y \le (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}, \tag{8.14}$$

$$\overline{h}(x),\ \overline{h}(y) \le \exp\left\{\cdot (2d)^{\exp O(r)} (2N)^{\log^* N \cdot \exp O(r)} \cdot (h+1)^3\right\}. \tag{8.15}$$

Proposition 8.7 will be proved in the next two sections. Now we use it to prove Theorem 3.5.

Proof of Theorem 3.5. Let $(x, y) \in \mathcal{C}$. Since $A \subseteq B$ we also have $(x, y) \in \mathcal{C}'$ where $B = A_0[f^{-1}, w]$, with f, w satisfying the conditions specified in Proposition 8.6. Then we use Proposition 8.7, to infer (8.14) and (8.15), and then we apply Lemma 6.4 to x and y, to show that x, y, x^{-1} and y^{-1} have representatives $\tilde{x}, \tilde{y}, \tilde{x}', \tilde{y}' \in \mathbb{Z}[X_1, \dots, X_r]$ with (3.16). □

8.3 Bounding the degree in Proposition 8.7

In this section we shall consider K as a function field in one variable, and we shall prove (8.14) using results of Brownawell and Masser [25] for function fields.

For the definition of valuations and height on function fields in one variable we refer to Section 7.2.

Proposition 8.8. *Let $\Bbbk$ be an algebraically closed field of characteristic 0, z a transcendental element over $\Bbbk$ and M a finite extension of $\Bbbk(z)$. Denote by $g_{M/\Bbbk}$ the genus of M and let S be a finite set of valuations of M. Denote by $\mathcal{O}_S$ the ring of S-integers of M, and by $\mathcal{O}_S^*$ its unit group. Consider the equation*

$$u_1 + \cdots + u_n = 0 \qquad \text{in} \qquad u_1, \ldots, u_n \in \mathcal{O}_S^*. \tag{8.16}$$

For every non-degenerate solution $u_1, \ldots, u_n$ of the above equation we have

$$H_M^*(u_1, \ldots, u_n) \leq \frac{1}{2}(n-1)(n-2)(|S| + 2g_{M/\Bbbk}).$$

Proof. This is in fact a variant of Corollary I of Brownawell and Masser [25], modified according to the remark after Theorem B of [25]. □

Proposition 8.9. *Let $\Bbbk$ be an algebraically closed field of characteristic 0, z a transcendental element over $\Bbbk$, M a finite extension of $\Bbbk(z)$, and $\overline{M}$ the algebraic closure of M. Denote by $g_{M/\Bbbk}$ the genus of M and let S be a finite set of valuations of M. Denote by $\mathcal{O}_S$ the ring of S-integers of M, and by $\mathcal{O}_S^*$ its unit group. Let $F(X,Y) = \sum_{(i,j)\in I} a_{ij} X^i Y^j \in \mathcal{O}_S[X,Y]$ with $a_{ij} \in \mathcal{O}_S^*$ for $(i,j) \in I$, be a polynomial which fulfils the condition that*

$$\begin{gathered} F \textbf{ is not divisible by } \text{any non-constant polynomial of the form} \\ X^m Y^n - \alpha \quad \text{or} \quad X^m - \alpha Y^n, \text{where } m, n \in \mathbb{Z}_{\geq 0} \text{ and } \alpha \in \overline{M}. \end{gathered} \tag{8.17}$$

Assume that $H_M(a_{ij}) \leq H_0$ for all $(i,j) \in I$. Then for every $x, y \in \mathcal{O}_S^*$ with

$$F(x,y) = 0$$

we have

$$H_M(x), H_M(y) \leq 2 \deg F \left(n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk} \right) + 2H_0 \right),$$

where $n(F)$ denotes the number of non-zero terms of F.

Proof. Since the coefficients of the polynomial F are S-units, we may consider the equation

$$\sum_{(i,j)\in I} a_{ij}x^iy^j = 0 \qquad \text{in} \quad x,y \in \mathcal{O}_S^* \tag{8.18}$$

as an equation of type (8.16). Let us fix a solution x, y of the equation. If there are vanishing sub-sums in the left hand side of (8.18) then all these vanishing sub-sums form individually an equation of type (8.16), and we get a system of the form

$$\begin{cases} \sum_{(i,j)\in I_1} a_{ij}x^iy^j = 0 & \text{in} \quad x,y \in \mathcal{O}_S^* \\ \cdots\cdots\cdots\cdots\cdots & \\ \sum_{(i,j)\in I_k} a_{ij}x^iy^j = 0 & \text{in} \quad x,y \in \mathcal{O}_S^*, \end{cases} \tag{8.19}$$

such that none of these equations has a proper vanishing subsum. Let $\mathcal{P} = (I_1, \ldots, I_k)$. As explained in Section 8.1 condition (8.17) implies that $\operatorname{rank}\Lambda(F,\mathcal{P}) = 2$. By dividing each equation of (8.19) by one of its terms we get

$$\begin{cases} 1 + \sum\limits_{(i,j)\in I_1\setminus\{(i_1,j_1)\}} \frac{a_{ij}}{a_{i_1j_1}} x^{i-i_1}y^{j-j_1} = 0 & \text{in} \quad x,y \in \mathcal{O}_S^* \\ \cdots\cdots\cdots\cdots\cdots & \\ 1 + \sum\limits_{(i,j)\in I_k\setminus\{(i_k,j_k)\}} \frac{a_{ij}}{a_{i_kj_k}} x^{i-i_k}y^{j-j_k} = 0 & \text{in} \quad x,y \in \mathcal{O}_S^*, \end{cases} \tag{8.20}$$

where we have $(i_l, j_l) \in I_l$ for $l = 1, \ldots, k$. Now we apply Proposition 8.8 to these equations. The number of terms of each equation is bounded above by $n(F)$, so we get

$$\begin{aligned} H_M\left(\frac{a_{ij}}{a_{i_lj_l}}x^{i-i_l}y^{j-j_l}\right) &\le H_M^*\left(\left(1, \frac{a_{ij}}{a_{i_lj_l}}x^{i-i_l}y^{j-j_l} \ : \ (i,j) \in I_l \setminus \{(i_l,j_l)\}\right)\right) \\ &\le H_M^*((a_{ij}x^iy^j \ : \ (i,j) \in I_l)) \le n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk}\right) \end{aligned}$$

for every $l = 1, \ldots, k$ and every $(i,j) \in I_l \setminus \{(i_l, j_l)\}$. Thus we have

$$\begin{aligned} H_M\left(x^{i-i_l}y^{j-j_l}\right) &\le n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk}\right) + H_M\left(\frac{a_{ij}}{a_{i_0j_0}}\right) \\ &\le n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk}\right) + 2H_0, \end{aligned}$$

which means that we have

$$H_M\left(x^ay^b\right) \le n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk}\right) + 2H_0,$$

for every $(a,b) = (u_1 - u_2, v_1 - v_2)$ with $(u_1, v_1), (u_2, v_2) \in I_l$ for some $l = 1, \ldots, k$. However $\Lambda(F, \mathcal{P}_{(x,y)}(F))$ is the $\mathbb{Z}$-module generated by these elements, and it has rank 2. Thus among

these generators there exist $(a_1, b_1), (a_2, b_2)$ with $a_1b_2 - a_2b_1 \neq 0$. By putting $z_1 := x^{a_1}y^{b_1}$ and $z_2 := x^{a_2}y^{b_2}$ we have

$$x^{a_1b_2-a_2b_1} = z_1^{b_2}z_2^{-b_1} \qquad y^{a_1b_2-a_2b_1} = z_2^{a_1}z_1^{-a_2},$$

and we get the estimate

$$\begin{aligned} H_M(x) &\leq \frac{1}{|a_1b_2 - a_2b_1|} H_M(z_1^{b_2}z_2^{-b_1}) \leq \frac{|b_2|H_M(z_1) + |b_1|H_M(z_2)}{|a_1b_2 - a_2b_1|} \\ &\leq 2 \deg F \left(n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk}\right) + 2H_0\right), \end{aligned}$$

and similarly

$$H_M(y) \leq 2 \deg F \left(n(F)^2 \cdot \left(|S| + 2g_{M/\Bbbk}\right) + 2H_0\right).$$

This concludes the proof of the proposition. □

Recall that $A = \mathbb{Z}[z_1, \ldots, z_r]$, K denotes the quotient field of A, $z_1, \ldots, z_q$ form a transcendence basis for K, $A_0 := \mathbb{Z}[z_1, \ldots, z_q]$, and $K_0 := \mathbb{Q}(z_1, \ldots, z_q)$. Further, let w be a primitive element of the extension K/K_0, which is integral over A_0 and has the properties specified in (i) of Proposition 8.6, and let $f \in A_0$ be an element with the properties specified in (ii) of Proposition 8.6. As above, put $B := A_0[w, f^{-1}]$.

Now let us fix $i \in \{1, \ldots, q\}$ and for each such fixed i put

$$\Bbbk_i := \mathbb{Q}(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_q).$$

Clearly, we have $A_0 \subseteq \overline{\Bbbk_i}[z_i]$, where $\overline{\Bbbk_i}$ denotes the algebraic closure of $\Bbbk_i$. Let $w^{(1)} = w, \ldots, w^{(D)}$ denote the conjugates of w over K_0, and put

$$\begin{aligned} M_i &:= \overline{\Bbbk_i}\left(z_i, w^{(1)}, \ldots, w^{(D)}\right), \\ B_i &:= \overline{\Bbbk_i}\left[z_i, w^{(1)}, \ldots, w^{(D)}, f^{-1}\right]. \end{aligned}$$

Then clearly M_i is the splitting field of the polynomial

$$\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X]$$

over $\overline{\Bbbk_i}[z_i]$, where $\mathcal{F}(X)$ is the minimal polynomial of w over K_0. Further, we have

$$B \subset B_i.$$

Let $\Delta_i := [M_i : \overline{\Bbbk}_i(z_i)]$ and denote by $g_{M_i/\overline{\Bbbk}_i}$ the genus of $M_i/\overline{\Bbbk}_i$, and by H_{M_i} the height taken with respect to $M_i/\overline{\Bbbk}_i$. In the following lemma we shall use the quantity

$$d_1 := \max(d_0, \deg f, \deg \mathcal{F}_1, \ldots, \deg \mathcal{F}_D), \tag{8.21}$$

and later we will use that by Proposition 8.6 we have the estimate

$$d_1 \le (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}. \tag{8.22}$$

Now we use Proposition 8.9 and Lemma 6.6 to prove statement (8.14) of Proposition 8.7:

Proof of (8.14). We denote by $w^{(1)} := w, \dots, w^{(D)}$ the conjugates of w over K_0, and for $\alpha \in K$ we denote by $\alpha^{(1)}, \dots, \alpha^{(D)}$ the conjugates of α corresponding to $w^{(1)}, \dots, w^{(D)}$. For $i = 1, \dots, n$ let $\Bbbk_i$, $\overline{\Bbbk}_i$, M_i, Δ_i have the same meaning as above. Let

$$S_i := \{v \in \mathcal{M}_{M_i} \ : \ v(z_i) < 0 \text{ or } v(f) > 0\}.$$

Since $w^{(j)} \in M_i$ and is integral over $\Bbbk_i[z_i]$, we have $w^{(j)} \in \mathcal{O}_{S_i}$ for $j = 1, \dots, D$. Since also $f^{-1} \in \mathcal{O}_{S_i}$ we have $\alpha^{(j)} \in \mathcal{O}_{S_i}$ for $\alpha \in B = A_0[f^{-1}, w]$, $j = 1, \dots, D$, $i = 1, \dots, q$.

Let $(x, y) \in \mathcal{C}'$. Then $x^{(j)}, y^{(j)}$ is a solution of the equation

$$F^{(j)}(x^{(j)}, y^{(j)}) = 0 \qquad \text{in} \quad x^{(j)}, y^{(j)} \in \mathcal{O}_{S_i}^*$$

for every $j = 1, \dots, D$, $i = 1, \dots, q$. Clearly the non-zero coefficients of $F^{(j)}(X, Y)$ are in $\mathcal{O}_{S_i}^*$, so by Proposition 8.9 we obtain that

$$\max(H_{M_i}(x^{(j)}), H_{M_i}(y^{(j)})) \le 2N \left(n(F)^2 \left(|S_i| + 2g_{M_i/\overline{\Bbbk}_i} \right) + 2H_0 \right), \tag{8.23}$$

where $H_0 := \max\limits_{i,j,u,v} H_{M_i}(a_{uv}^{(j)})$. By $\deg \tilde{a}_{uv} \le d$ and Lemma 6.3 we have $\overline{\deg}\, a_{uv} \le (2d)^{\exp O(r)}$, which together with Lemma 6.7 gives

$$H_0 \le \Delta_i (2d)^{\exp O(r)}. \tag{8.24}$$

Now we have to estimate the genus of $M_i/\overline{\Bbbk}_i$ and the cardinality of S_i. First, using Lemma 7.6 for $\overline{\Bbbk}_i[z_i]$ and the polynomial $\mathcal{F}(X) = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D$, in view of the bounds in (i) of Proposition 8.6 we get

$$g_{M_i/\overline{\Bbbk}_i} \le \Delta_i D \max_{1 \le k \le D} \deg_{z_i} \mathcal{F}_k \le \Delta_i D (2d_0)^{\exp O(r)} \le \Delta_i (2d)^{\exp O(r)}. \tag{8.25}$$

To bound $|S_i|$ we mention that every valuation of $\overline{\Bbbk}_i(z_i)$ can be extended to at most $[M_i : \overline{\Bbbk}_i(z_i)] = \Delta_i$ valuations of M_i. Thus the number of valuations v of M_i with $v(z_i) < 0$ is bounded by Δ_i and similarly, the number of valuations v of M_i with $v(f) > 0$ is bounded above by $\Delta_i \deg_{z_i} f$. Hence altogether we have

$$\begin{aligned} |S_i| &\le \Delta_i + \Delta_i \deg_{z_i} f \le \Delta_i (1 + \deg f) \\ &\le \Delta_i (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}, \end{aligned} \tag{8.26}$$

where in the estimates we have used (ii) of Proposition 8.6.

Now turning again our attention to the estimate (8.23), and using (8.25) and (8.26) we get

$$\max(H_{M_i}(x^{(j)}), H_{M_i}(y^{(j)})) \le \Delta_i (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}. \tag{8.27}$$

Now it is the time to use Lemma 6.6, which together with (8.27), $D \le d^r$, $q \le r$ and (8.22) proves that

$$\begin{aligned}\overline{\deg}\, x, \overline{\deg}\, y &\le qDd_1 + \sum_{i=1}^{q} \Delta_i^{-1} \sum_{j=1}^{D} H_{M_i}(x^{(j)}) \\ &\le (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}.\end{aligned}$$

This concludes the proof of (8.14) of Proposition 8.7. □

8.4 Bounding the height in Proposition 8.7

8.4.1 The result for the number field case

For the definition of the absolute logarithmic height of algebraic numbers and of the height of polynomials we refer to Section 2.1.

In this section we present a version of Theorem 2.1 of [11]. Let Γ be a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^2$. Let $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$ be a basis of Γ modulo Γ_{tors}. Put

$$h_w := \max\big(1, h(\mathbf{w}_1), \dots, h(\mathbf{w}_r)\big).$$

Denote by K the smallest number field such that $\Gamma \subset (K^*)^2$, and put $d := [K : \mathbb{Q}]$. Let S be the minimal finite set of places of K containing all the infinite places of K and having the property that $\Gamma \subset (\mathcal{O}_S^*)^2$ and denote by s the cardinality of S. Recall that $\mathbf{P}$ is defined in (2.5). The discriminant of the field K is denoted by D_K.

Let $f(X, Y) \in \overline{\mathbb{Q}}[X, Y]$ be a polynomial which is not divisible by any non-constant polynomial of the shape $aX^m Y^n - b$ or $aX^m - bY^n$ for some $a, b \in \overline{\mathbb{Q}}$, $m, n \in \mathbb{Z}_{\ge 0}$. We mention that in this case f is also not divisible by any polynomial which depends on exactly one of the variables X, Y, since then it would be divisible by a polynomial of the shape $aX - b$ or $aY - b$, respectively. Put $N := \deg f$ for the total degree of f. Let L be the field extension of K generated by the coefficients of f. Put

$$\begin{aligned}&\delta := \deg_X f + \deg_Y f, \quad H := \max(1, h(f)), \\ &C_0 := \big(e^{13}\delta^7 d^3 r\big)^{r+3} s \cdot \frac{\mathbf{P}^{2\delta^2}}{\log \mathbf{P}} h_w^r \cdot \log^* \big(\max(\delta d s \mathbf{P}, \delta h_w)\big).. \\ &C_0^* := \big(\delta \cdot d \cdot s \cdot \log \mathbf{P} \cdot D_K (\log^* D_K)^{d-1}\big)^{O(s^2)} \cdot \mathbf{P}^{2\delta^2}.\end{aligned}$$

Let $\mathcal{C} \subset (\overline{\mathbb{Q}}^*)^2$ be the curve defined by $f(x,y) = 0$.

Proposition 8.10. *Assume that f is absolutely irreducible. Then for every point $\mathbf{x} = (x,y) \in \mathcal{C} \cap \Gamma$ we have*

$$h(x) + h(y) \leq C_0 H.$$

Proof. This is just Theorem 2.2 with a slightly rewritten bound. □

Proposition 8.11. *Assume that $\Gamma = \mathcal{O}_S^*$ but not that f is absolutely irreducible. Then for every point $\mathbf{x} = (x,y) \in \mathcal{C} \cap \Gamma$ we have*

$$h(x) + h(y) \leq C_0^*(H + 2N).$$

Proof. This is a weaker version of Proposition 8.10. We shortly explain how this result is deduced from Proposition 8.10. If $f(x,y) = 0$ then there exists an absolutely irreducible factor $g(X,Y)$ of f, which then fulfils the conditions of Proposition 8.10, thus we may apply the latter to g. Further, since g divides f it is also well known that $h(g) \leq h(f) + 2N$ (see Proposition B.7.3 of [43]).

We also have to take care of the dependence on h_w and r, more precisely to estimate h_w and r in the case $\Gamma = \mathcal{O}_S^*$. If we take $\Gamma := (\mathcal{O}_S^*)^2$ then one can bound the number of generators r of Γ by $2s-2$ and we may choose a system of fundamental S-units to get a set of generators for $(\mathcal{O}_S^*)^2$, so that the height of these elements in this fundamental system is bounded. More precisely by Lemma 2 of [41] we can choose the generators such that

$$h_0 \leq c_1 R_S,$$

where $c_1 := 29e\sqrt{s-2}d^{s-1}(\log^* d) \cdot ((s-1)!)^2/(2^{s-2}d^{s-1})$, and R_S is the S-regulator of K. For the S-regulator by using Lemma 3 of [27] and Lemma 2.1 of [9] (for the original result see Louboutin [53]) we can derive the bound

$$R_S \leq |D_K|^{\frac{1}{2}}(\log^* |D_K|)^{d-1} \cdot (\log P)^s.$$

Combining these estimates the bound of our proposition follows by a simple computation. We mention that a much sharper bound could have been deduced, but this estimate is more than enough for our purpose. □

8.4.2 Specializations

In this section we shall use many specializations which map K to a number field, in order to be able to profit from our results from Section 8.4.1. The main feature of these specializations, called Győry-Kronecker specializations is that using sufficiently many of them, there will be at least one, which makes possible to extend effective results over number fields to similar results over finitely generated domains. Such specializations were first used by Győry [39] and [40], however, here we introduce and use the refined version of this specialization method due to Evertse and Győry [32].

First for every $\mathbf{u} \in \mathbb{Z}^q$ we may replace z_i by u_i for $i = 1, \ldots, q$. This defines a homomorphism from a subring of K_0 to $\mathbb{Q}$. More precisely, for fixed $\mathbf{u} \in \mathbb{Z}^q$ we consider the map

$$\varphi_{\mathbf{u}} : \ \alpha \mapsto \alpha(\mathbf{u}) = \frac{g_1(\mathbf{u})}{g_2(\mathbf{u})}$$

which is defined for every $\alpha = \frac{g_1}{g_2} \in K_0$ with $g_1, g_2 \in A_0$ and with the additional property $g_2(\mathbf{u}) \neq 0$, and has its image in $\overline{\mathbb{Q}}$. Now we wish to extend this to a ring homomorphism from B to $\overline{\mathbb{Q}}$. Thus we will impose some restrictions on $\mathbf{u}$. Recall that $K = K_0(w)$, $B = A_0[f^{-1}, w]$ with $f \in A_0$ and that $\mathcal{F}$ is the minimal polynomial of w, and $f \in A_0$, both with the properties specified in Proposition 8.6. Let $\Delta_{\mathcal{F}}$ denote the discriminant of $\mathcal{F}$ with the convention $\Delta_{\mathcal{F}} = 1$ if $\mathcal{F}$ is a linear polynomial. Put

$$\mathcal{H} := \Delta_{\mathcal{F}} \cdot \mathcal{F}_D \cdot f,$$

observe that $\mathcal{H} \in A_0$ and assume that $\mathbf{u}$ is chosen such that $\mathcal{H}(\mathbf{u}) \neq 0$. Put

$$\begin{cases} d_0^* = \max(\deg \mathcal{F}_1, \ldots, \deg \mathcal{F}_D) \\ h_0^* = \max(h(\mathcal{F}_1), \ldots, h(\mathcal{F}_D)) \end{cases} \qquad \begin{cases} d_1^* = \max(d_0^*, \deg f) \\ h_1^* = \max(h_0^*, h(f)). \end{cases}$$

By Proposition 8.6 we infer that

$$\begin{cases} d_0^* \leq (2d_0)^{\exp O(r)} \leq (2d)^{\exp O(r)} \\ h_0^* \leq (2d_0)^{\exp O(r)}(h_0 + 1) \leq (2d)^{\exp O(r)}(h + 1) \\ d_1^* \leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} \\ h_1^* \leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} \cdot (h + 1). \end{cases} \tag{8.28}$$

Thus we clearly have

$$\deg \mathcal{H} \leq (2D - 2) \cdot d_0^* + d_0^* + d_1^* \leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}. \tag{8.29}$$

Now let $\mathbf{u} \in \mathbb{Z}^q$ be fixed such that $\mathcal{H}(\mathbf{u}) \neq 0$. Thus the polynomial

$$\mathcal{F}_{\mathbf{u}} := X^D + \mathcal{F}_1(\mathbf{u})X^{D-1} + \cdots + \mathcal{F}_D(\mathbf{u})$$

has non-zero discriminant, and since $\mathcal{F}_D(\mathbf{u}) \neq 0$ it has D distinct non-zero roots. Let us denote these numbers by $w^{(1)}(\mathbf{u}), \ldots, w^{(D)}(\mathbf{u})$.

To extend our map $\varphi_{\mathbf{u}}$ to B we use the representation (8.12) of elements $\alpha \in B$. Namely, for each $j = 1, \ldots, D$ we may define the function $\varphi_{\mathbf{u},j}$ such that for $\alpha \in B$ written as

$$\alpha = \sum_{i=1}^{D-1} (P_i/Q)\, w^i, \qquad \text{where } P_0, \ldots, P_{D-1}, Q \in A_0,\ \gcd(P_0, \ldots, P_{D-1}, Q) = 1, \tag{8.30}$$

we set

$$\varphi_{\mathbf{u},j}(\alpha) := \sum_{i=1}^{D-1} (P_i(\mathbf{u})/Q(\mathbf{u})) \left(w^{(j)}(\mathbf{u})\right)^i. \tag{8.31}$$

This is well-defined, since for $\alpha \in B$ the polynomial Q must divide a power of f, hence $Q(\mathbf{u}) \neq 0$. By this we described exactly D ways to extend $\varphi_{\mathbf{u}}$ from K_0 to B. Clearly, the map $\varphi_{\mathbf{u},j}$ defined above is a ring homomorphism from B to $\overline{\mathbb{Q}}$, thus any unit of B is mapped to a non-zero element of $\overline{\mathbb{Q}}$ by any of the specializations defined above. Put

$$K_{\mathbf{u},j} := \mathbb{Q}(w^{(j)}(\mathbf{u})) \qquad \text{for} \qquad j = 1, \ldots, D, \tag{8.32}$$

and denote by $\Delta_{K_{\mathbf{u},j}}$ the discriminant of the algebraic number field $K_{\mathbf{u},j}$.

We recall that Lemma 6.8 provides an upper bound for $|\Delta_{K_{\mathbf{u},j}}|$, Lemma 6.9 bounds the height of $\alpha^{(j)}(\mathbf{u})$ for $\mathbf{u} \in \mathbb{Z}^q$ in terms of the size of $\alpha \in B$ and some parameters of B and Lemma 6.10 shows that if we take a sufficiently large number of specializations, then there is at least one specialization among them (say corresponding to $\mathbf{u} \in \mathbb{Z}^q$), such that $\overline{h}(\alpha)$ for $\alpha \in B$ can be bounded by the heights of the images of α by the specializations $\varphi_{\mathbf{u},j}$ for $j = 1, \ldots, D$.

8.4.3 Conclusion of the proof of Proposition 8.7

In this subsection we combine the specialization method and the result for the number field case presented in subsections 8.4.1 and 8.4.2, in order to prove (8.15).

Proof of (8.15) of Proposition 8.7. Since in the case $q = 0$ we are in the number field case our Theorem 2.1 of [11] will give a much better bound than stated in Proposition

8.7. So we may restrict ourselves to the case $q > 0$. Let $\mathcal{P}$ be a fixed partition of I and $(x, y) \in \mathcal{C}'$ be a fixed solution associated with $\mathcal{P}$. Choose $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$ and $k \in \{1, \dots, D\}$, and consider the corresponding specialization $\varphi_{\mathbf{u},k}$ defined in (8.31), where later we shall specify some further requirements on $\mathbf{u}$ and k when we shall apply Lemma 6.10. Then we have the notation

$$\begin{aligned} &\varphi_{\mathbf{u},k}(x) = x^{(k)}(\mathbf{u}), \qquad \varphi_{\mathbf{u},k}(y) = y^{(k)}(\mathbf{u}), \\ &\varphi_{\mathbf{u},k}(a_{ij}) = a_{ij}^{(k)}(\mathbf{u}) \qquad \text{for} \quad (i,j) \in I. \end{aligned} \tag{8.33}$$

Put $F_{\mathbf{u},k}(X, Y) := \sum_{(i,j)\in I} a_{ij}^{(k)}(\mathbf{u}) X^i Y^j$, let $K_{\mathbf{u},k}$ be the field defined in (8.32), $S_{\mathbf{u},k}$ the set of places of $K_{\mathbf{u},k}$ containing all infinite places and those finite places which lie above prime ideals dividing $f(\mathbf{u})$. Since we clearly have

$$\varphi_{\mathbf{u},k}(B) \subseteq \mathcal{O}_{S_{\mathbf{u},k}},$$

from $(x, y) \in \mathcal{C}'$ we get

$$F_{\mathbf{u},k}\left(x^{(k)}(\mathbf{u}), y^{(k)}(\mathbf{u})\right) = 0 \qquad \text{in} \quad x^{(k)}(\mathbf{u}), y^{(k)}(\mathbf{u}) \in \mathcal{O}_{S_{\mathbf{u},k}}^*. \tag{8.34}$$

Now we shall apply Lemma 6.10. Since in the previous section we have proved (8.14), i.e.

$$\overline{\deg}\, x, \overline{\deg}\, y \leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)},$$

now in view of (8.28) we may apply Lemma 6.10 with some

$$N_0 \leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)}$$

to infer that the set

$$\mathcal{S} := \{\mathbf{u} \in \mathbb{Z}^q \ : \ |\mathbf{u}| \leq N_0, \ \mathcal{H}(\mathbf{u}) \neq 0\}$$

is non-empty. Taking also (8.28) into account we have

$$\begin{aligned} \overline{h}(x) &\leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} (h+1)^2 H_x, \\ \overline{h}(y) &\leq (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} (h+1)^2 H_y, \end{aligned} \tag{8.35}$$

where $H_x := \max\{h(x^{(k)}(\mathbf{u})) \ : \ \mathbf{u} \in \mathcal{S}, \ k = 1, \dots, D\}$ and $H_y := \max\{h(y^{(k)}(\mathbf{u})) \ : \ \mathbf{u} \in \mathcal{S}, \ k = 1, \dots, D\}$.

To finish the proof, the last step is to estimate H_x and H_y using Proposition 8.11 for equation (8.34). We fix any $\mathbf{u} \in \mathcal{S}$ and $k = 1, \dots, D$. By Lemma 6.8 and in view of (8.28) we get that

$$\begin{aligned} |\Delta_{K_{\mathbf{u},k}}| &\leq D^{2D-1} \left((d_0^*)^q e^{h_0^*} \max(1, |\mathbf{u}|^{d_0^*})\right)^{2D-2} \\ &\leq \exp\left\{(2d)^{\exp O(r)} \cdot (h+1) \cdot (\log^* N)^2\right\}, \end{aligned} \tag{8.36}$$

and $[K_{\mathbf{u},k} : \mathbb{Q}] \le D$.

To estimate $h(F_{\mathbf{u},k})$ we bound the height of its coefficients, i.e. $h(a_{ij}^{(k)}(\mathbf{u}))$ for $(i,j) \in I$. For this we use first Lemma 6.3, which in view of $\deg \tilde{a}_{ij} < d$ and $h(\tilde{a}_{ij}) < h$ gives

$$\overline{\deg}\, a_{ij} \le (2d)^{\exp O(r)} \qquad \overline{h}(a_{ij}) \le (2d)^{\exp O(r)}(h+1).$$

This together with Lemma 6.9 gives for every $(i,j) \in I$ the estimate

$$h\left(a_{ij}^{(k)}(\mathbf{u})\right) \le (\log^* N)^2 (2d)^{\exp O(r)}(h+1),$$

which in turn proves

$$h(F_{\mathbf{u},k}) \le n(F) \cdot \max h\left(a_{ij}^{(k)}(\mathbf{u})\right) \le N^2 (\log^* N)^2 (2d)^{\exp O(r)}(h+1). \tag{8.37}$$

We also have to estimate the cardinality of $S_{K_{\mathbf{u},k}}$. For this, we first bound the absolute value of $f(\mathbf{u})$ by the elementary computation

$$\begin{aligned} |f(\mathbf{u})| &\le (\deg f)^q \cdot e^{h(f)} \cdot (\max(1,|\mathbf{u}|))^{\deg f} \le (d_1^*)^q \cdot e^{h_1^*} \cdot (\max(1,|\mathbf{u}|))^{d_1^*} \\ &\le \exp\left\{(2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} \cdot (h+1)\right\}. \end{aligned}$$

Clearly we have $s := |S_{K_{\mathbf{u},j}}| \le D(1+\omega(f(\mathbf{u})))$, where $\omega(f(\mathbf{u}))$ denotes the number of distinct prime factors of $f(\mathbf{u})$. Thus we get

$$\begin{aligned} s &\le O\left(d^r \log^* |f(\mathbf{u})| / \log^* \log^* |f(\mathbf{u})|\right) \\ &\le (2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} \cdot (h+1). \end{aligned} \tag{8.38}$$

Further, for the maximum of the norm of the prime ideals belonging to $S_{K_{\mathbf{u},k}}$ we have the estimate

$$\mathbf{P} \le |f(\mathbf{u})|^D \le \exp\left\{(2d)^{\exp O(r)} \cdot (2N)^{\log^* N \cdot \exp O(r)} \cdot (h+1)\right\}. \tag{8.39}$$

Now we shall show that for the polynomial $F_{\mathbf{u},l}$ we have

$$\begin{aligned} &F_{\mathbf{u},l} \textbf{ is not divisible by} \text{ any non-constant polynomial of the form} \\ &\qquad X^m Y^n - \alpha \quad \text{or} \quad X^m - \alpha Y^n, \text{where } m,n \in \mathbb{Z}_{\ge 0} \text{ and } \alpha \in \overline{K}_{\mathbf{u},l}. \end{aligned} \tag{8.40}$$

The coefficients a_{ij} of F are units in B, thus all these coefficients are mapped to non-zero elements $a_{ij}^{(l)}(\mathbf{u})$ by the specialization $\varphi_{\mathbf{u},l}$, so the partitions of the polynomial F are just the same as the partitions of the polynomial $F_{\mathbf{u},l}$. If $\operatorname{rank} \Lambda(F,\mathcal{P}) = 2$ then we also have $\operatorname{rank} \Lambda(F_{\mathbf{u},l},\mathcal{P}) = 2$. Further, if $\operatorname{rank} \Lambda(F,\mathcal{P}) = 1$ then we also have $\operatorname{rank} \Lambda(F_{\mathbf{u},l},\mathcal{P}) = 1$ and by Proposition 8.1 the corresponding system of polynomials $g_1, \dots, g_k$ (see Section 8.1) has

the property $\gcd(g_1, \dots, g_k) = 1$ in $K[X]$. Thus there exist polynomials $u_1, \dots, u_k \in A[X]$ and a constant $R \in A$ with

$$u_1 g_1 + \cdots + u_k g_k = R, \tag{8.41}$$

and by Proposition 8.3 we see that R can be chosen such that it has a representative $\tilde{R}$ with

$$\deg \tilde{R} \le d(4N)^{\log_2^* N}, \qquad h(\tilde{R}) \le (4N)^{\log_2^* N + 2}(d+1)rh.$$

This R fulfils all assumptions made for R in Proposition 8.6, so assume that f and B have been chosen in Proposition 8.6 such that $R \in B^*$. Now we apply the specialization $\varphi_{\mathbf{u},l}$ to (8.41) to infer that

$$(u_1)_{\mathbf{u},l}(g_1)_{\mathbf{u},l} + \cdots + (u_k)_{\mathbf{u},l}(g_k)_{\mathbf{u},l} = R_{\mathbf{u}}^{(l)},$$

where for a polynomial P with coefficients in A, $(P)_{\mathbf{u},l}$ denotes the polynomial obtained by applying $\varphi_{\mathbf{u},l}$ the the coefficients of P. Since $R \in B^*$ we have $R_{\mathbf{u}}^{(l)} \neq 0$ hence $\gcd((g_1)_{\mathbf{u},l}, \dots, (g_k)_{\mathbf{u},l}) = 1$ in $K_{\mathbf{u},l}$. By Proposition 8.1 this proves (8.40). So the polynomial $F_{\mathbf{u},l}$ cannot have any non-constant factor of the shape $aX^mY^n - b$ or $aX^m - bY^n$ for some $a, b \in \overline{\mathbb{Q}}$, $m, n \in \mathbb{Z}_{\ge 0}$. Thus the solution set of equation (8.34) fulfills the conditions of Proposition 8.11, so combining this by statements (8.37), (8.38), (8.39), (8.36) and $[K_{\mathbf{u},k} : \mathbb{Q}] \le D$ we get the estimate

$$h(x^{(k)}(\mathbf{u})), h(y^{(k)}(\mathbf{u})) \le \exp\left\{(2d)^{\exp O(r)}(2N)^{\log^* N \cdot \exp O(r)} \cdot (h+1)^3\right\},$$

for every $\mathbf{u} \in \mathcal{S}$ and $k = 1, \dots, D$, which provides the same upper bound for H_x and H_y. Now combining this latter estimate with (8.35) we get the desired bound (8.15). This concludes the proof of Proposition 8.7. □

Chapter 9

Proof of the results from Section 3.4

9.1 A reduction

In this section we reduce Theorem 3.6 to two propositions and using a result of Everste and Győry [32] we show how Theorem 3.6 can be deduced from these propositions.

Proposition 9.1. *Let A be a finitely generated domain as above, $\overline{\Gamma}$ the above-defined division group and $F(X,Y) \in A[X,Y]$ a polynomial which fulfils the condition (3.22). Then there exists a suitably large effectively computable constant C_3 such that for every $(x,y) \in \mathcal{C}$ there exists an exponent*

$$m_0 < N^6(2d)^{\exp\{C_3(r+s)\}}(h+1)^{4s}$$

for which we have

$$x^{m_0} \in \Gamma, \qquad y^{m_0} \in \Gamma.$$

We remark that in this proposition the value of the exponent m_0, although bounded by (9.1), it may depend on the choice of the pair $(x,y) \in \mathcal{C}$. In contrast, in the statement (i) of Theorem 3.6 the exponent m is uniform, i.e. it does not depend on the pair $(x,y) \in \mathcal{C}$. Now we deduce statement (i) of Theorem 3.6 from the above Proposition 9.1.

Proof of Theorem 3.6 (i). Let C_3 be the constant specified in Proposition 9.1 and define

$$M_0 := \left[N^6(2d)^{\exp\{C_3(r+s)\}}(h+1)^{4s}\right].$$

Put

$$m := \operatorname{lcm}(1,\ldots,M_0).$$

Then by Proposition 9.1 we clearly have

$$x^m, y^m \in \Gamma$$

for every $(x, y) \in \mathcal{C}$.

Using the estimate

$$\pi(M) \le \frac{4}{3}\frac{M}{\log M}$$

of Rosser and Schönfeld [64] for the number $\pi(M)$ of primes up to M we get

$$\begin{aligned}\operatorname{lcm}(1, \dots, M) &\le \prod_{p \le M} p^{[\log M/\log p]} \le \prod_{p \le M} p^{\log M/\log p} \\ &= \prod_{p \le M} M \le M^{\pi(M)} \le M^{\frac{4}{3}\frac{M}{\log M}} \le e^{\frac{4}{3}M}.\end{aligned}$$

Thus we have the estimate

$$m \le \exp\left\{N^6 (2d)^{\exp O(r+s)} (h+1)^{4s}\right\},$$

which concludes the proof of (i) of Theorem 3.6. □

Next let us fix m to be the integer specified in (i) of Theorem 3.6 and consider the set

$$\mathcal{C}_1 := \left\{(x_0, y_0) \in \Gamma^2 \mid \exists\, x, y \in \overline{\Gamma} : x^m = x_0, y^m = y_0, F(x, y) = 0\right\}. \tag{9.1}$$

Proposition 9.2. *Let $(x_0, y_0) \in \mathcal{C}_1$. Then there exist representatives $\tilde{x}_0$ and $\tilde{y}_0$ for x_0 and y_0, respectively, with the property*

$$\begin{aligned}&\deg \tilde{x}_0, \deg \tilde{y}_0 \le \exp\left\{N^6 (2d)^{\exp O(r+s)} (h+1)^{4s}\right\} \\ &h(\tilde{x}_0), h(\tilde{y}_0) \le \exp\left\{\exp\left\{N^{12} (2d)^{\exp O(r+s)} (h+1)^{8s}\right\}\right\}\end{aligned} \tag{9.2}$$

To deduce (ii) of Theorem 3.6 from the above proposition we need the following result of Evertse and Győry [32].

Lemma 9.3. *Let $\gamma_1, \dots, \gamma_s \in K^*$ be multiplicatively independent elements, and assume that for $\gamma_0 \in K^*$ we have*

$$\gamma_0 = \gamma_1^{k_1} \dots \gamma_s^{k_s}.$$

Further, assume that for $i = 0, \dots, s$ we have pairs of representatives (g_{i1}, g_{i2}) for γ_i such that

$$\begin{cases} \deg f_1, \dots, \deg f_t, \deg g_{0,1}, \deg g_{0,2}, \dots, \deg g_{s,1}, \deg g_{s,2} \le d_2 \\ h(f_1), \dots, h(f_t), h(g_{0,1}), h(g_{0,2}), \dots, h(g_{s,1}), h(g_{s,2}) \le h_2, \end{cases} \tag{9.3}$$

for some real numbers $d_2, h_2 > 1$. Then we also have

$$|k_i| \le (2d_2)^{\exp O(r+s)} (h_2+1)^{2s}, \qquad \text{for } i = 1, \dots, s. \tag{9.4}$$

Proof. This is Corollary 7.3 of Evertse and Győry [32]. □

Proof of (ii) of Theorem 3.6. Let m be the exponent specified in (i) of Theorem 3.6. Then by $x^m \in \Gamma$ and $y^m \in \Gamma$ we have

$$x^m = \gamma_1^{t_{1,x}} \dots \gamma_s^{t_{s,x}}, \qquad y^m = \gamma_1^{t_{1,y}} \dots \gamma_s^{t_{s,y}} \tag{9.5}$$

with certain integer exponents $t_{1,x}, \dots, t_{s,x}$ and $t_{1,y}, \dots, t_{s,y}$. Now by our assumption on $\gamma_1, \dots, \gamma_s$ and by Proposition 9.2 we see that x^m, y^m, $\gamma_1, \dots, \gamma_s$ have representatives with degrees and heights below the bound

$$\exp\left\{\exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}\right\},$$

which together with Lemma 9.3 applied to the relations (9.5) concludes the proof of statement (ii) of Theorem 3.6. □

9.2 Proof of Proposition 9.1

We split the proof of Proposition 9.1 into several steps, each being presented in a separate subsection:

- for $(x, y) \in \mathcal{C}$ we bound the degree of the field $K(x, y)$ over K;
- we estimate the smallest positive integer exponent M such that for $(x, y) \in \mathcal{C}$ we have $x^M, y^M \in \Gamma_K$, where Γ_K denotes the K-closure of Γ, i.e. the largest subgroup of $\overline{\Gamma}$ which belongs to K^*;
- for $\gamma \in \Gamma_K$ we estimate the smallest positive integer exponent $m(\gamma)$ such that $\gamma^{m(\gamma)} \in \Gamma$;
- we conclude the proof of Proposition 9.1.

We also mention that using Lemma 6.3 and (3.23) we have the estimates

$$\overline{\deg}\,\gamma_i \le (2d)^{\exp(O(r))}, \qquad \overline{h}(\gamma_i) \le (2d)^{\exp(O(r))}(h+1) \tag{9.6}$$

for $i = 1 \dots, s$, and

$$\overline{\deg}\,a_{ij} \le (2d)^{\exp(O(r))}, \qquad \overline{h}(a_{ij}) \le (2d)^{\exp(O(r))}(h+1). \tag{9.7}$$

for $(i, j) \in I$. These estimates will be frequently used in the rest of the Chapter.

9.2.1 Bounding the degree of $K(x, y)$

Let $(x, y) \in \mathcal{C}$. We shall give a bound on the degree of the field $L := K(x, y)$ over K. Since $x, y \in \overline{\Gamma}$, there exist $m_x, m_y \in \mathbb{Z}_{>0}$ such that $x^{m_x}, y^{m_y} \in \Gamma$. Take the least common multiple of m_x and m_y and denote it by m_{xy}. Then $x^{m_{xy}}, y^{m_{xy}} \in \Gamma \subset K$, so we have

$$[K(x, y) : K] \leq m_{xy}.$$

In order to estimate m_{xy} put $F_{x,y}(X, Y) := F(xX, yY)$. Then for any embedding $\sigma : L \hookrightarrow \overline{K}$, we have $F(\sigma(x), \sigma(y)) = 0$, hence

$$F_{x,y}\left(\frac{\sigma(x)}{x}, \frac{\sigma(y)}{y}\right) = 0.$$

Since $x^{m_x} \in K$ we also have $\sigma(x)^{m_x} \in K$, hence

$$\left(\frac{\sigma(x)}{x}\right)^{m_{xy}} = 1,$$

and similarly,

$$\left(\frac{\sigma(y)}{y}\right)^{m_{xy}} = 1.$$

Thus the distinct embeddings $\sigma : K(x, y) \hookrightarrow \overline{K}$ give rise to distinct solutions $\left(\frac{\sigma(x)}{x}, \frac{\sigma(y)}{y}\right)$ of the equation

$$F_{x,y}(\rho_1, \rho_2) = 0 \qquad \text{in} \qquad \rho_1, \rho_2 \quad \text{roots of unity.}$$

However, by the result of Beukers and Smyth [13] this equation has at most $22(\deg F)^2$ solutions in roots of unity. Thus we have proved that

$$[K(x, y) : K] \leq 22(\deg F)^2 = 22N^2. \tag{9.8}$$

9.2.2 Bounding the exponent M

Let again $x, y \in \mathcal{C}$. Let us denote by Γ_K the K-closure of Γ, i.e. the largest subgroup of $\overline{\Gamma}$ belonging to K^*. Then we have $\Gamma \subseteq \Gamma_K \subseteq \overline{\Gamma}$. Now we shall give an upper bound for the minimal exponent M such that

$$x^M, y^M \in \Gamma_K.$$

Let $d_x := [K(x) : K]$. Then we have $d_x \leq 22N^2$. By $x \in \overline{\Gamma}$ there exists $m_x \in \mathbb{N}$ with $x^{m_x} \in \Gamma$. Put $\gamma := x^{m_x}$. Then the minimal polynomial $f_x(X)$ of x over K divides

$$X^{m_x} - \gamma = \prod_{i=1}^{m_x} (X - \rho^i x),$$

where ρ is a primitive root of unity of order m_x. Thus

$$f_x(X) = \prod_{j=1}^{d_x} \left(X - \rho^{i_j} x\right) \in K[X]$$

with suitable distinct choices of $i_j \in \{1, \ldots, m_x\}$. Hence there exists a root of unity ρ such that

$$\rho x^{d_x} \in K^*.$$

Now let us estimate the order l of ρ. Clearly, $\rho \in K(x)$. Further, since $[K : K_0] \leq d^{r-q}$ (see Proposition 6.1 (i)) we have $[K(x) : K_0] \leq d_x \cdot d^{r-q}$, and by $\rho \in K(x)$ this gives the estimate $[K(\rho) : K_0] \leq d_x \cdot d^{r-q}$, which gives $[K_0(\rho) : K_0] \leq d_x \cdot d^{r-q}$. However, since $K_0 = \mathbb{Q}(z_1, \ldots, z_q)$, with algebraically independent elements $z_1, \ldots, z_q$, we have $[K_0(\rho) : K_0] = [\mathbb{Q}(\rho) : \mathbb{Q}]$, hence

$$[\mathbb{Q}(\rho) : \mathbb{Q}] \leq d_x \cdot d^{r-q}.$$

On the other hand ρ is a root of unity of order l, so

$$[\mathbb{Q}(\rho) : \mathbb{Q}] = \varphi(l),$$

which gives the estimate

$$\varphi(l) \leq d_x \cdot d^{r-q}.$$

Now using the estimate

$$\varphi(l) \gg \frac{l}{\log \log l}$$

of Rosser and Schönfeld [64] we infer

$$l \ll \left(d_x d^{r-q}\right)^2,$$

where the constants implied by $\ll$ and $\gg$ are absolute constants. However, since $\rho x^{d_x} \in K^*$ we have $(\rho x^{d_x})^l \in K^*$ hence $x^{d_x \cdot l} \in K^*$, which by $d_x \leq 22N^2$ proves that

$$M \leq d_x \cdot l \ll d_x^3 d^{2(r-q)} \ll N^6 d^{2(r-q)}. \tag{9.9}$$

9.2.3 Bounding the exponent $m(\gamma)$

The next step is to take an arbitrary element

$$\gamma \in \Gamma_K \setminus \Gamma.$$

Since $\gamma \in \Gamma_K \subseteq \overline{\Gamma}$ there exists a minimal natural number $m(\gamma)$ such that

$$\gamma^{m(\gamma)} \in \Gamma.$$

We now estimate $m(\gamma)$. Clearly, for such an $m(\gamma)$ we have

$$\gamma^{m(\gamma)} = \gamma_1^{t_1} \dots \gamma_s^{t_s} \tag{9.10}$$

and without loss of generality we may suppose that $0 \le t_i < m(\gamma)$ for $i = 1, \dots, s$. Indeed, if we take v_i with $v_i \equiv t_i \pmod{m(\gamma)}$ and $0 \le v_i < m(\gamma)$ for $i = 1, \dots, s$, then considering $\gamma' := \gamma_1^{v_1} \dots \gamma_s^{v_s}$ we have $m(\gamma) = m(\gamma')$, and for bounding $m(\gamma)$ we may just replace γ by γ'. So we start with the relation

$$\gamma^{m(\gamma)} = \gamma_1^{t_1} \dots \gamma_s^{t_s}, \qquad 0 \le t_i < m(\gamma). \tag{9.11}$$

Now we first bound $\overline{\deg}\,\gamma$ and $\overline{h}(\gamma)$.

9.2.3.1 Bounding $\overline{\deg}\,\gamma$

Recall that the elements $\gamma_1, \dots, \gamma_s \in K^*$ are given by corresponding representation pairs $(g_1, h_1), \dots, (g_s, h_s)$, which fulfil (3.23). First we extend the domain A to a larger domain B such that the "numerators" and "denominators" of $\gamma_1, \dots, \gamma_s$ are all units of B. More precisely, let $\gamma_{i1} := g_i(z_1, \dots, z_r)$ and $\gamma_{i2} := h_i(z_1, \dots, z_r)$ for $i = 1, \dots, s$. Then we have the following:

Proposition 9.4. *There exists a non-zero $f \in A_0$ such that*

$$A \subseteq A_0[w, f^{-1}] =: B, \quad \gamma_{i1}, \gamma_{i2} \in A_0[w, f^{-1}]^* \quad \textit{for } i = 1, \dots, s \tag{9.12}$$

and

$$\deg f \le (2s+1)(2d)^{\exp O(r)}, \qquad h(f) \le (2s+1)(2d)^{\exp O(r)}(h+1). \tag{9.13}$$

Proof. This is a simple consequence of (ii) of Proposition 6.1. Indeed, we have $k = 2s$, and in view of (3.23) we may take $d^{**} = d$ and $h^{**} = h$. □

We use the notation of Section 6.2.

By (6.15), (3.23), (6.3) and (9.13) we have the estimate

$$d_1 \le (2s+1)(2d)^{\exp O(r)}. \tag{9.14}$$

By Lemma 6.7 we have

$$\max_{i,j} H_{M_i}\left(\gamma_k^{(i)}\right) \le \Delta_i \left(2D\overline{\deg}\,\gamma_k + (2d_0)^{\exp O(r)}\right) \le \Delta_i (2d)^{\exp O(r)}. \tag{9.15}$$

Further, by (9.11) we have

$$mH_{M_i}(\gamma^{(j)}) \le \sum_{k=1}^{s} t_k H_{M_i}(\gamma_k^{(j)})$$

which means

$$H_{M_i}(\gamma^{(j)}) \le \sum_{k=1}^{s} \frac{t_k}{m} H_{M_i}(\gamma_k^{(j)}) \le \sum_{k=1}^{s} H_{M_i}(\gamma_k^{(j)}) \le s\Delta_i(2d)^{\exp O(r)}.$$

Thus by Lemma 6.6 we get

$$\begin{aligned} \overline{\deg}\,\gamma &\le qDd_1 + \sum_{i=1}^{q} \Delta_i^{-1} \sum_{j=1}^{D} H_{M_i}(\gamma^{(j)}) \\ &\le qD(2s+1)(2d)^{\exp O(r)} + \sum_{i=1}^{q} \Delta_i^{-1} Ds\Delta_i(2d)^{\exp O(r)} \\ &\le s(2d)^{\exp O(r)}. \end{aligned} \tag{9.16}$$

9.2.3.2 Bounding $\overline{h}(\gamma)$

We shall use the notation of Section 6.3. Let $\varphi_{\mathbf{u},j}$ be a specialization map on the domain B as defined in Section 9.2.3.1. Then applying $\varphi_{\mathbf{u},j}$ to the relation (9.11) we get

$$\gamma^{(j)}(\mathbf{u})^{m(\gamma)} = \gamma_1^{(j)}(\mathbf{u})^{t_1} \dots \gamma_s^{(j)}(\mathbf{u})^{t_s}, \tag{9.17}$$

which gives

$$m(\gamma)h(\gamma^{(j)}(\mathbf{u})) \le \sum_{i=1}^{s} t_i h(\gamma_i^{(j)}(\mathbf{u}))$$

leading to the estimate

$$h(\gamma^{(j)}(\mathbf{u})) \le \sum_{i=1}^{s} \frac{t_i}{m(\gamma)} h(\gamma_i^{(j)}(\mathbf{u})) \le \sum_{i=1}^{s} h(\gamma_i^{(j)}(\mathbf{u})). \tag{9.18}$$

By Lemma 6.9 and (9.6) we have

$$h(\gamma_i^{(j)}(\mathbf{u})) \le (2d)^{\exp(O(r))} \left(h + 1 + \log\max(1, |\mathbf{u}|)\right). \tag{9.19}$$

Using the domain B specified in Section 9.2.3.1 by (9.12) we have

$$d_1^* \le (2s+1)(2d)^{\exp(O(r))}, \qquad h_1^* \le (2s+1)(2d)^{\exp(O(r))}(h+1). \tag{9.20}$$

Now in order to use Lemma 6.10 we may choose $N_0 := (2s+1)(2d)^{\exp(O(r))}$ providing the bound

$$h(\gamma_i^{(j)}(\mathbf{u})) \le (2d)^{\exp(O(r))}(h+1)s,$$

which together with (9.18) gives

$$h(\gamma^{(j)}(\mathbf{u})) \le s^2(2d)^{\exp(O(r))}(h+1),$$

and the use of Lemma 6.10 leads to the estimate

$$\overline{h}(\gamma) \le s^6(2d)^{\exp(O(r))}(h+1)^2. \tag{9.21}$$

9.2.3.3 Bounding the exponents in (9.11)

Lemma 9.5. *Let $\gamma_0, \gamma_1, \dots, \gamma_s \in K^*$ be multiplicatively dependent elements, and assume that for $i = 0, \dots, s$ we have*

$$\overline{\deg}\,\gamma_i \le (2d)^{\exp O(r+s)}, \qquad \overline{h}(\gamma_i) \le (2d)^{\exp O(r+s)}(h+1)^2. \tag{9.22}$$

Then there exist integers $k_0, \dots, k_s$ not all equal to 0 such that

$$\gamma_0^{k_0} \dots \gamma_s^{k_s} = 1$$

and

$$|k_i| \le (2d)^{\exp O(r+s)}(h+1)^{2s}, \qquad \text{for } i = 0, \dots, s. \tag{9.23}$$

Proof. This is a variant of Lemma 7.2 of Evertse and Győry in [32]. To prove this result it would be necessary to redo the long proof of Lemma 7.2 of [32], with most part of it completely unchanged. So here we only indicate those points which should be changed in the proof of Lemma 7.2 of [32] to get our Lemma 9.5. The first point is, that after defining the elements $\gamma_{\mathbf{v}}$ we have to estimate their $\overline{\deg}$, i.e. we get

$$\overline{\deg}\,\gamma_{\mathbf{v}} \le \sum_{i=0}^{s} v_i \overline{\deg}\,\gamma_i \le V \cdot (2d)^{\exp(O(r+s))}.$$

Thus using our Proposition 6.2 we get the same estimate

$$\deg f \le V^{\exp O(r+s)}$$

as in [32]. From this point we have to redo identically the computation of the proof of Lemma 7.2 of [32], just with the bounds (9.22) instead of the bounds given in the proof of Evertse and Győry for $\overline{\deg}\,\gamma_i$ and $\overline{h}(\gamma_i)$, and finally we get the estimate (9.23). □

Now applying Lemma 9.5 to our identity (9.11) we get the desired bound

$$|m(\gamma)|, |t_i| \le (2d)^{\exp(O(r+s))}(h+1)^{2s}. \tag{9.24}$$

9.2.4 Concluding the proof of Proposition 9.1

In Section 9.2.2 we proved that for a given $(x, y) \in \mathcal{C}$ there exists an exponent M with (9.9) such that

$$x^M, y^M \in \Gamma_K.$$

Further, by Section 9.2.3 there exist exponents $m(x^M)$ and $m(y^M)$ with

$$m(x^M), m(y^M) \leq (2d)^{\exp(O(r+s))}(h+1)^{2s},$$

such that

$$(x^M)^{m(x^M)}, (y^M)^{m(y^M)} \in \Gamma.$$

Put $m_0 := M \cdot m(x^M) \cdot m(y^M)$. Then we have

$$m_0 \leq N^6(2d)^{\exp(O(r+s))}(h+1)^{4s}.$$

Denoting by C_3 the constant implied by the $O(\cdot)$ symbol in the last inequality the proof of Proposition 9.1 is concluded.

9.3 Proof of Proposition 9.2

9.3.1 Bounding the degree

We shall use the notation of Section 6.1 however we shall extend our domain A to a larger domain B in a different way than we did in Section 9.2. More precisely, we choose f and thus B as described in the following proposition:

Proposition 9.6. *There exists a non-zero $f \in A_0$ with*

$$\begin{aligned} &\deg f \leq sN^2(2d)^{\exp O(r)}, \\ &h(f) \leq sN^2(2d)^{\exp O(r)}(h+1), \end{aligned} \tag{9.25}$$

such that with $B := A_0[f^{-1}, w]$

$$\begin{aligned} &A \subseteq B, \\ &\gamma_{i1}, \gamma_{i2} \in B^* \quad \textit{for } i = 1, \ldots, s, \\ &a_{ij} \in B^* \quad \textit{for } (i,j) \in I. \end{aligned} \tag{9.26}$$

Proof. This is a simple consequence of (ii) of Proposition 6.1. Indeed, we have $k := 2s + |I| \leq O(sN^2)$, and in view of (3.23) we may take $d^{**} = d$ and $h^{**} = h$. □

Put $B := A_0[w, f^{-1}]$. Then we clearly have $A \subseteq B \subseteq K$. Recall that for a fixed $i \in \{1, \dots, q\}$ in Section 6.2 we introduced the notation

$$\begin{aligned} \Bbbk_i &:= \mathbb{Q}(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_q), \\ \overline{\Bbbk}_i &:= \text{ the algebraic closure of } \Bbbk_i, \\ M_i &:= \overline{\Bbbk}_i(z_i, w^{(1)}, \dots, w^{(D)}), \\ B_i &:= \overline{\Bbbk}_i[z_i, f^{-1}, w^{(1)}, \dots, w^{(D)}], \end{aligned} \tag{9.27}$$

where w is the element specified in Proposition 6.1 and $w^{(1)}, \dots, w^{(D)}$ denote the conjugates of w over K_0. Further, we used the notation

$$\begin{aligned} \Delta_i &:= [M_i : \overline{\Bbbk}_i(z_i)], \\ g_{M_i} &:= \text{ the genus of } M_i/\overline{\Bbbk}_i, \\ H_{M_i} &:= \text{ the height with respect to } M_i/\overline{\Bbbk}_i. \end{aligned} \tag{9.28}$$

Let $\mathcal{M}_{M_i}$ denote the set of places of M_i and define

$$S_i := \{v \in \mathcal{M}_{M_i} \mid v(z_i) < 0 \quad \text{or} \quad v(f) > 0\}. \tag{9.29}$$

Then we have the estimates

$$\begin{aligned} |S_i| &\le \Delta_i + \Delta_i \deg_{z_i} f \le \Delta_i(1 + \deg f) \\ &\le \Delta_i s N^2 (2d)^{\exp O(r)}, \end{aligned} \tag{9.30}$$

and

$$g_{M_i} \le \Delta_i D \max_{1 \le k \le D} \deg_{z_i} \mathcal{F}_k \le \Delta_i D (2d_0)^{\exp O(r)} \le \Delta_i (2d)^{\exp O(r)}. \tag{9.31}$$

Now let x, y be such that $F(x, y) = 0$ and $x^m, y^m \in \Gamma$ and put $\tilde{M}_i := M_i(x, y)$. Denote by $\tilde{S}_i$ the set of places of M_i lying above places of S_i and denote by $g_{\tilde{M}_i}$ the genus of the extension $\tilde{M}_i/\Bbbk_i$. By (9.8) we clearly have

$$[\tilde{M}_i : M_i] \le 22N^2.$$

Now we wish to bound $g_{\tilde{M}_i}$ and $|\tilde{S}_i|$. By the Riemann-Hurwitz formula we have

$$2g_{\tilde{M}_i} - 2 = [\tilde{M}_i : M_i] \cdot (2g_{M_i} - 2) + \sum_{v \in \mathcal{M}_{M_i}} \sum_{\substack{V \mid v \\ V \in \mathcal{M}_{\tilde{M}_i}}} (e(V \mid v) - 1).$$

The valuations $v \notin S_i$ do not ramify, i.e. $e(V \mid v) = 1$ for V above v, hence

$$\begin{aligned} 2g_{\tilde{M}_i} - 2 &= [\tilde{M}_i : M_i] \cdot (2g_{M_i} - 2) + \sum_{v \in S_i} \sum_{\substack{V \mid v \\ V \in \tilde{S}_i}} (e(V \mid v) - 1) \\ &= [\tilde{M}_i : M_i] \cdot (2g_{M_i} - 2) + \sum_{v \in S_i} [\tilde{M}_i : M_i] - |\tilde{S}_i| \\ &= [\tilde{M}_i : M_i] \cdot (2g_{M_i} - 2) + |S_i|[\tilde{M}_i : M_i] - |\tilde{S}_i|. \end{aligned}$$

This gives us

$$2g_{\tilde{M}_i} + |\tilde{S}_i| = [\tilde{M}_i : M_i] \cdot (2g_{M_i} - 2 + |S_i|) + 2,$$

which provides at the same time the upper bounds for $g_{\tilde{M}_i}$ and $|\tilde{S}_i|$ given below:

$$g_{\tilde{M}_i} \leq 22N^2(2g_{M_i} + |S_i|) \leq \Delta_i s N^4 (2d)^{\exp O(r)}, \tag{9.32}$$

and similarly,

$$|\tilde{S}_i| \leq \Delta_i s N^4 (2d)^{\exp O(r)}. \tag{9.33}$$

We will use Proposition 8.9 to bound $H_{\tilde{M}_i}(x)$ and $H_{\tilde{M}_i}(y)$ by

$$H_{\tilde{M}_i}(x), H_{\tilde{M}_i}(y) \leq 2N \left[(N+1)^4 (|\tilde{S}_i| + g_{\tilde{M}_i}) + 2H_0 \right], \tag{9.34}$$

where H_0 is an upper bound for $H_{\tilde{M}_i}(a_{uv}^{(j)})$. By Lemma 6.7 and by (9.7) we get

$$\begin{aligned} H_{\tilde{M}_i}(a_{uv}^{(j)}) &\leq [\tilde{M}_i : \overline{\Bbbk}_i(z_i)] \cdot \left(2D \overline{\deg}\, a_{uv} + (2d_0)^{\exp O(r)} \right) \\ &\leq 22N^2 \Delta_i (2d)^{\exp O(r)} \leq N^2 \Delta_i (2d)^{\exp O(r)} =: H_0 \end{aligned}$$

This together with (9.34) provides the estimate

$$H_{\tilde{M}_i}(x), H_{\tilde{M}_i}(y) \leq \Delta_i s N^9 (2d)^{\exp O(r)}, \tag{9.35}$$

which together with (3.25) proves

$$H_{\tilde{M}_i}(x_0) = H_{\tilde{M}_i}(x^m) \leq m H_{\tilde{M}_i}(x) \leq \Delta_i \exp\left\{ N^6 (2d)^{\exp O(r+s)} (h+1)^{4s} \right\},$$

and the same bound for $H_{\tilde{M}_i}(y_0)$. Now using $H_{\tilde{M}_i}(x_0) = [\tilde{M}_i : M_i] \cdot H_{M_i}(x_0)$ and $H_{\tilde{M}_i}(y_0) = [\tilde{M}_i : M_i] \cdot H_{M_i}(y_0)$ we obtain

$$H_{M_i}(x_0), H_{M_i}(y_0) \leq \Delta_i \exp\left\{ N^6 (2d)^{\exp O(r+s)} (h+1)^{4s} \right\},$$

which together with Lemma 6.6 proves the desired estimate

$$\overline{\deg}\, x_0, \overline{\deg}\, y_0 \leq \exp\left\{ N^6 (2d)^{\exp O(r+s)} (h+1)^{4s} \right\}. \tag{9.36}$$

9.3.2 Preparations for bounding the height

Lemma 9.7. *Let $\mathcal{R}$ be an integral domain, $H(X) := \sum_{i=0}^{n} c_i X^i \in \mathcal{R}[X]$ be a polynomial, and ρ a primitive m^{th} root of unity. Then we have*

$$\prod_{j=1}^{m} H(\rho^j X) \in \mathcal{R}[X^m].$$

Proof. There is no loss of generality, to assume that $\mathcal{R} = \mathbb{Z}[c_0, \dots, c_n]$, where $c_0, \dots, c_n$ are independent variables. Let $K_c := \mathbb{Q}(c_0, \dots, c_n)$ and $\overline{K}_c$ the algebraic closure of K_c. We may write

$$H(X) = c_n \prod_{k=1}^{n} (X - \alpha_k),$$

where $\alpha_1, \dots, \alpha_n \in \overline{K}_c$. Thus

$$\begin{aligned}\prod_{j=0}^{m-1} H(\rho^j X) &= c_n^m \prod_{k=1}^{n} \prod_{j=0}^{m-1} (\rho^j X - \alpha_k) = c_n^m \rho^{nm(m-1)/2} \prod_{k=1}^{n} (X^m - \rho^{-j} \alpha_k^m) \\ &= c_n^m (-1)^{n(m-1)} \prod_{k=1}^{n} (X^m - \alpha_k^m).\end{aligned}$$

Since the coefficients of $\prod_{k=1}^{n} (X^m - \alpha_k^m)$ are elementary symmetric polynomials with integral coefficients in $\alpha_1, \dots, \alpha_n$, they are in fact elements of $\mathbb{Z}[c_0/c_n, \dots, c_{n-1}/c_n]$. Hence $\prod_{j=0}^{m-1} H(\rho^j X) =: G(X^m)$ with $G \in K_c[X]$. But the coefficients of G are integral over $\mathcal{R}$, and $\mathcal{R}$ is integrally closed closed, hence they belong to $\mathcal{R}$. □

Proposition 9.8. *Let ρ be a primitive m^{th} root of unity. Then there exists a polynomial $G(U, V) = \sum_{(i,j) \in \mathcal{J}} b_{ij} U^i V^j \in A[U, V]$ with $b_{ij} \neq 0$ for $(i, j) \in \mathcal{J}$, such that*

$$G(X^m, Y^m) = \prod_{k=0}^{m-1} \prod_{l=0}^{m-1} F(\rho^k X, \rho^l Y), \tag{9.37}$$

and such that the coefficients b_{ij} of G have representatives $\tilde{b}_{ij}$ with

$$\deg \tilde{b}_{ij} \leq m^2 d, \qquad h(\tilde{b}_{ij}) \leq m^2 (h + 2 \log(N + 1)) \tag{9.38}$$

Proof. Put $R_0 := \mathbb{Z}[a_{pq} \ : \ (p, q) \in I]$, where we adjoin all coefficients a_{pq}, $(p, q) \in I$ of F to $\mathbb{Z}$. First we consider the polynomial $H(X, Y) := \prod_{l=0}^{m-1} F(X, \rho^l Y)$ as a polynomial in one variable (namely in Y) over the integral domain $R := R_0[X]$. Then by Lemma 9.7 we see that $H(X, Y) \in R[Y^m]$. Using again Lemma 9.7 for the polynomial

$$\prod_{k=0}^{m-1} \prod_{l=0}^{m-1} F(\rho^k X, \rho^l Y) = \prod_{k=0}^{m-1} H(\rho^k X, Y),$$

and the ring $R_1 := R_0[Y^m]$ we infer that

$$\prod_{k=0}^{m-1} \prod_{l=0}^{m-1} F(\rho^k X, \rho^l Y) \in R_0[X^m, Y^m],$$

so we clearly have $\prod_{k=0}^{m-1}\prod_{l=0}^{m-1} F(\rho^k X, \rho^l Y) \in A[X^m, Y^m]$, thus the existence of a polynomial $G(U,V) \in A[U,V]$ with (9.37) is proved.

Now we have to prove the estimates for the coefficients of G. Recall that $F(X,Y) = \sum_{(p,q)\in I} a_{pq}X^pY^q$, and by assumption (see (3.23)) we are given representatives $\tilde{a}_{pq}$ such that

$$\deg \tilde{a}_{pq} \le d, \qquad h(\tilde{a}_{pq}) \le h.$$

Put

$$\tilde{F}_{kl} := \tilde{F}(\rho^k X, \rho^l Y).$$

For a polynomial F with complex coefficients let us denote by $||F||_1$ the one-norm of F, i.e. the sum of the absolute values of the coefficients of F.

Then we have

$$||\tilde{F}_{kl}||_1 = ||\tilde{F}||_1 = \sum_{(p,q)\in I} ||\tilde{a}_{pq}||_1$$

and

$$||\tilde{G}||_1 \le \prod_{k=0}^{m-1}\prod_{l=0}^{m-1} ||\tilde{F}_{kl}||_1 \le ||\tilde{F}||_1^{m^2}.$$

This shows that

$$h(\tilde{G}) \le m^2 \log||F||_1 \le m^2(h(\tilde{F}) + \log|I|) \le m^2(h + 2\log(N+1)),$$

and this proves

$$h(\tilde{b}_{ij}) \le m^2(h + 2\log(N+1)).$$

Further, the coefficient b_{kl} is a sum of products of the terms a_{pq}, each summand consisting of at most m^2 multiplicands, so we have

$$\deg \tilde{b}_{ij} \le m^2 d.$$

□

Lemma 9.9. *Let $G(X,Y)$ be the polynomial defined in Proposition 9.8. Then $G(X,Y)$ is divisible by a non-constant polynomial of the form $X^aY^b - \alpha$ or $X^a - \alpha Y^b$ with $\alpha \in \overline{K}^*$, $a,b \in \mathbb{Z}_{\ge 0}$ if and only if $F(X,Y)$ is divisible by a non-constant polynomial of the form $X^uY^v - \beta$ or $X^u - \beta Y^v$ with $\beta \in \overline{K}^*$, $u,v \in \mathbb{Z}_{\ge 0}$.*

Proof. Clearly we may assume $\gcd(a,b) = 1$, otherwise we factorize $X^aY^b - \alpha$ or $X^a - \alpha Y^b$ and we get a similar factor of G with the property $(a,b) = 1$. Then $X^{ma}Y^{mb} - \alpha$ or $X^{ma} - \alpha Y^{mb}$ divides $G(X^m, Y^m)$, which also means that $X^aY^b - \alpha'$ or $X^a - \alpha' Y^b$ divides

$G(X^m, Y^m)$ with a suitable $\alpha' \in \overline{K}^*$. However, by $\gcd(a, b) = 1$ we know that $X^aY^b - \alpha'$ or $X^a - \alpha' Y^b$ is absolutely irreducible, so if it divides $G(X^m, Y^m)$ then it divides one of $F_{kl}(X, Y)$, but this means that $\rho^{-ka} X \rho^{-lb} Y - \alpha'$ or $\rho^{-ka} X - \alpha' \rho^{-lb} Y$ divides $F(X, Y)$. The converse is trivial, so this concludes the proof of the Lemma. □

Lemma 9.10. *The set $\mathcal{C}_1$ defined in (9.1) is equal to the set*

$$\left\{(x_0, y_0) \in \Gamma^2 \mid G(x_0, y_0) = 0\right\}. \tag{9.39}$$

Proof. Denote the set in (9.39) by $\mathcal{C}_2$. If $(x_0, y_0) \in \mathcal{C}_1$ then there exist $x, y \in \overline{\Gamma}$ with $x^m = x_0$, $y^m = y_0$ such that $F(x, y) = 0$. So clearly

$$0 = \prod_{k=0}^{m-1} \prod_{l=0}^{m-1} F(\rho^k x, \rho^l y) = G(x^m, y^m) = G(x_0, y_0),$$

thus $(x_0, y_0) \in \mathcal{C}_2$.

Conversely, if $(x_0, y_0) \in \mathcal{C}_2$ then we have $G(x_0, y_0) = 0$. All the m^{th} roots of x_0 and y_0 are zeros of the polynomials

$$X^m - x_0 = \prod_{k=0}^{m-1} (X - \rho^k x_0^{1/m})$$

and

$$X^m - y_0 = \prod_{l=0}^{m-1} (X - \rho^l y_0^{1/m}),$$

respectively, with any fixed choice $x_0^{1/m}$ and $y_0^{1/m}$ of an m^{th} roots of x_0 and y_0. Then we have

$$\prod_{k=0}^{m-1} \prod_{l=0}^{m-1} F_{kl}\left(x_0^{1/m}, y_0^{1/m}\right) = G\left((x_0^{1/m})^m, (y_0^{1/m})^m\right) = 0,$$

so there exist $k, l \in \{0, \ldots, m-1\}$ with

$$F_{kl}\left(x_0^{1/m}, y_0^{1/m}\right) = 0,$$

i.e. we have

$$F\left(\rho^k x_0^{1/m}, \rho^l y_0^{1/m}\right) = 0.$$

Thus by the choice $x = \rho^k x_0^{1/m}$ and $y = \rho^l y_0^{1/m}$ there exist $x, y \in \overline{\Gamma}$ with $x^m = x_0$, $y^m = y_0$ and $F(x, y) = 0$, however these conditions just mean that $(x_0, y_0) \in \mathcal{C}_1$. This concludes the proof of our Lemma. □

9.3.3 Bounding the height of elements of $\mathcal{C}_1$

Now we will use the specialization method to bound $\overline{h}(x_0), \overline{h}(y_0)$ for any $(x_0, y_0) \in \mathcal{C}_1$. More precisely, we prove

Proposition 9.11. *If $(x_0, y_0) \in \mathcal{C}_1$ then we have*

$$\overline{\deg}\, x_0, \overline{\deg}\, y_0 \le \exp\left\{N^6 (2d)^{\exp O(r+s)} (h+1)^{4s}\right\} \tag{9.40}$$

$$\overline{h}(x_0), \overline{h}(y_0) \le \exp\left\{\exp\left\{N^{12} (2d)^{\exp O(r+s)} (h+1)^{8s}\right\}\right\} \tag{9.41}$$

This sub-section is devoted to the proof of Proposition 9.11. Recall that the coefficients b_{ij} of G have representatives $\tilde{b}_{ij}$ with

$$\deg \tilde{b}_{ij} \le m^2 d, \qquad h(\tilde{b}_{ij}) \le m^2 (h + 2\log(N+1)).$$

Thus by Lemma 6.3 and by (3.25) we have

$$\overline{\deg}\, b_{ij}, \overline{h}(b_{ij}) \le \exp\left\{N^6 (2d)^{\exp O(r+s)} (h+1)^{4s}\right\}. \tag{9.42}$$

Again, we have to extend the domain A to a larger domain. For this, we use a suitable version of Proposition 6.2. More precisely we have

Proposition 9.12. *Let $R \in A$ be an arbitrary non-zero element having a representative $\tilde{R}$ with the property*

$$\deg \tilde{R}, h(\tilde{R}) \le \exp\left\{N^{12} (2d)^{\exp O(r+s)} (h+1)^{8s}\right\}. \tag{9.43}$$

Then there exists a non-zero $f \in A_0$ such that for $B := A_0[w, f^{-1}]$

$$\begin{aligned} &A \subseteq B, \\ &\gamma_{i1}, \gamma_{i2} \in B^* \quad \text{for } i = 1, \dots, s, \\ &b_{ij} \in B^* \quad \text{for } (i,j) \in \mathcal{J}, \\ &R \in B^*. \end{aligned} \tag{9.44}$$

Further, f can be chosen such that it fulfils

$$\deg f, h(f) \le \exp\left\{N^{12} (2d)^{\exp O(r+s)} (h+1)^{8s}\right\}. \tag{9.45}$$

Proof. This is a variant of Proposition 6.2. We clearly know that the $\overline{\deg}$ and $\overline{h}$ of the elements $\gamma_{i1}, \gamma_{i2} \in A_0[w, f^{-1}]^*$ for $i = 1, \dots, s$, $b_{ij} \in A_0[w, f^{-1}]^*$ for $(i,j) \in \mathcal{J}$, and $R \in A_0[w, f^{-1}]^*$ are bounded by

$$\exp\left\{N^{12} (2d)^{\exp O(r+s)} (h+1)^{8s}\right\}.$$

thus by Proposition 6.2 the estimate (9.45) follows at once. □

The element $R \in A$ in the above Proposition will be specified later during the proof, and will be chosen such that it fulfils condition (9.43).

Proof of Proposition 9.11. For the case $q = 0$ we are in the number field case, and for this case much better bounds are provided by [11], so we may assume $q > 0$.

Estimate (9.40) is the same as (9.36). Now we prove the estimate (9.41) using the specialization method described in Section 6.3.

Let $\mathcal{P}$ be a fixed partition of $\mathcal{J}$ and $(x_0, y_0) \in \mathcal{C}_2$ be a fixed solution associated with $\mathcal{P}$.

Put $B := A_0[w, f^{-1}]$. Then for the quantities $d_0^*, d_1^*, h_0^*, h_1^*$ defined in (6.21) we have the following estimates:

$$\begin{aligned} &d_0^* \le (2d)^{\exp O(r)}, \\ &h_0^* \le (2d)^{\exp O(r)}(h+1), \\ &d_1^*, h_1^* \le \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}. \end{aligned} \tag{9.46}$$

Thus for $\mathcal{H} := \Delta_{\mathcal{F}} \cdot \mathcal{F}_D \cdot f \in A_0$ we have

$$\deg \mathcal{H} \le (2D-2)d_0^* + d_0^* + d_1^* \le \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}.$$

We shall choose $\mathbf{u} \in \mathbb{Z}^q$ such that $\mathcal{H}(\mathbf{u}) \neq 0$ and consider the extended specializations $\varphi_{\mathbf{u},k}$ defined in (6.24). Then we have

$$\begin{aligned} &\varphi_{\mathbf{u},k}(x_0) = x_0^{(k)}(\mathbf{u}), \qquad \varphi_{\mathbf{u},k}(y_0) = y_0^{(k)}(\mathbf{u}), \\ &\varphi_{\mathbf{u},k}(b_{ij}) = b_{ij}^{(k)}(\mathbf{u}) \qquad \text{for} \quad (i,j) \in I. \end{aligned} \tag{9.47}$$

Later we shall specify some further requirements on $\mathbf{u}$ and k to be able to apply Lemma 6.10.

The polynomial $G(X, Y)$ is mapped by the extended specialization $\varphi_{\mathbf{u},k}$ to the polynomial $G_{\mathbf{u},k}(X, Y) := \sum_{(i,j)\in\mathcal{J}} b_{ij}^{(k)}(\mathbf{u}) X^i Y^j$. Let $K_{\mathbf{u},k}$ be the field defined in (6.28), and $S_{\mathbf{u},k}$ the set of places of $K_{\mathbf{u},k}$ containing all infinite places and those finite places which lie above prime ideals dividing $f(\mathbf{u})$. Since we clearly have

$$\varphi_{\mathbf{u},k}(\Gamma) \subseteq \varphi_{\mathbf{u},k}(B^*) \subseteq \mathcal{O}_{S_{\mathbf{u},k}}^*,$$

from $(x_0, y_0) \in \mathcal{C}_2$ we get

$$G_{\mathbf{u},k}\left(x^{(k)}(\mathbf{u}), y^{(k)}(\mathbf{u})\right) = 0 \qquad \text{in} \quad x_0^{(k)}(\mathbf{u}), y_0^{(k)}(\mathbf{u}) \in \mathcal{O}_{S_{\mathbf{u},k}}^*. \tag{9.48}$$

The next step of the proof is to apply Lemma 6.10. By (9.40) and in view of (9.46) in Lemma 6.10 we may choose

$$N_0 \le \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}$$

to infer that the set

$$\mathcal{S} := \{\mathbf{u} \in \mathbb{Z}^q \ : \ |\mathbf{u}| \le N_0, \ \mathcal{H}(\mathbf{u}) \neq 0\}$$

is non-empty. Put

$$H_1 := \max\{h(x_0^{(k)}(\mathbf{u})), h(y_0^{(k)}(\mathbf{u})) \ : \ \mathbf{u} \in \mathcal{S}, \ k = 1, \dots, D\}.$$

Then using (9.46) and Lemma 6.10 we infer

$$\overline{h}(x_0), \overline{h}(y_0) \le \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\} H_1. \tag{9.49}$$

The last step is to estimate H_1. Fix any $\mathbf{u} \in \mathcal{S}$ and $k = 1, \dots, D$. First using Lemma 6.8 and (9.46) we can estimate the parameters of the field $K_{\mathbf{u},k}$:

$$\begin{aligned} |\Delta_{K_{\mathbf{u},k}}| &\le D^{2D-1}\left((d_0^*)^q e^{h_0^*}\max(1, |\mathbf{u}|^{d_0^*})\right)^{2D-2} \\ &\le \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}, \end{aligned} \tag{9.50}$$

and $[K_{\mathbf{u},k} : \mathbb{Q}] \le D$.

To bound $h(G_{\mathbf{u},k})$ we first have to estimate the height of its coefficients. Lemma 6.9 together with (9.42) gives

$$h\left(b_{ij}^{(k)}(\mathbf{u})\right) \le \exp\left\{N^6(2d)^{\exp O(r+s)}(h+1)^{4s}\right\}.$$

This leads to the estimate

$$\begin{aligned} h(G_{\mathbf{u},k}) &\le n(G) \cdot \max h\left(b_{ij}^{(k)}(\mathbf{u})\right) \\ &\le \exp\left\{N^6(2d)^{\exp O(r+s)}(h+1)^{4s}\right\}. \end{aligned} \tag{9.51}$$

To bound the cardinality of $S_{K_{\mathbf{u},k}}$ first we estimate

$$\begin{aligned} |f(\mathbf{u})| &\le (\deg f)^q \cdot e^{h(f)} \cdot (\max(1, |\mathbf{u}|))^{\deg f} \le (d_1^*)^q \cdot e^{h_1^*} \cdot (\max(1, |\mathbf{u}|))^{d_1^*} \\ &\le \exp\left\{\exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}\right\}. \end{aligned}$$

Since $s := |S_{K_{\mathbf{u},j}}| \le D(1 + \omega(f(\mathbf{u})))$, where $\omega(f(\mathbf{u}))$ denotes the number of distinct prime factors of $f(\mathbf{u})$, we get

$$\begin{aligned} s &\le O\left(d^r \log^* |f(\mathbf{u})| / \log^* \log^* |f(\mathbf{u})|\right) \\ &\le \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}. \end{aligned} \tag{9.52}$$

For the maximum of the norm of the prime ideals belonging to $S_{K_{\mathbf{u},k}}$ we have

$$\mathbf{P} \le |f(\mathbf{u})|^D \le \exp\left\{\exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}\right\}. \tag{9.53}$$

Now in order to be able to use Proposition 8.11 for the equation (9.48) we have to prove that the polynomial $G_{\mathbf{u},k}$ fulfils the condition

$$\begin{gathered} G_{\mathbf{u},l} \textbf{ is not divisible by} \text{ any non-constant polynomial of the form} \\ X^m Y^n - \alpha \quad \text{or} \quad X^m - \alpha Y^n, \text{where } m, n \in \mathbb{Z}_{\geq 0} \text{ and } \alpha \in \overline{K}_{\mathbf{u},l}. \end{gathered} \tag{9.54}$$

The coefficients b_{ij} of G are units in B, thus all these coefficients are mapped to non-zero elements $b_{ij}^{(l)}(\mathbf{u})$ by the specialization $\varphi_{\mathbf{u},l}$. Thus the partitions of the polynomial G are just the same as the partitions of the polynomial $G_{\mathbf{u},l}$.

If $r(G, \mathcal{P}) = 2$ then we also have $r(G_{\mathbf{u},k}, \mathcal{P}) = 2$. Further, if $r(G, \mathcal{P}) = 1$ then we also have $r(G_{\mathbf{u},k}, \mathcal{P}) = 1$, and by Proposition 8.1 the corresponding system of polynomials $g_1, \ldots, g_k$ (see Section 8.1.1) has the property $\gcd(g_1, \ldots, g_k) = 1$ in $K[X]$. Thus there exist polynomials $u_1, \ldots, u_k$ and a non-zero constant $R \in A$ with

$$u_1 g_1 + \cdots + u_k g_k = R, \tag{9.55}$$

and by Corollary 8.5 we see that R can be chosen such that it has a representative $\tilde{R}$ with

$$\deg \tilde{R}, h(\tilde{R}) \leq \exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}.$$

Assume that $f \in A_0$ and the domain B are chosen in Proposition 9.12 such that $R \in B^*$. Now apply the specialization $\varphi_{\mathbf{u},k}$ to the relation (9.55) to infer that

$$(u_1)_{\mathbf{u},k}(g_1)_{\mathbf{u},k} + \cdots + (u_k)_{\mathbf{u},k}(g_k)_{\mathbf{u},k} = R_{\mathbf{u}}^{(k)}.$$

Since $R \in B^*$ we have $R_{\mathbf{u}}^{(k)} \neq 0$, and we see that $\gcd((g_1)_{\mathbf{u},k}, \ldots, (g_k)_{\mathbf{u},k}) = 1$ in $K_{\mathbf{u},k}[X]$. However, the above argument by Proposition 8.1 implies that $G_{\mathbf{u},k}$ fulfils (9.54).

So the polynomial $G_{\mathbf{u},k}$ cannot have any non-constant factor of the shape $aX^m Y^n - b$ or $aX^m - bY^n$ for some $a, b \in \overline{\mathbb{Q}}$, $m, n \in \mathbb{Z}_{\geq 0}$. Thus the solution set of equation (9.48) fulfills the conditions of Proposition 8.11, so combining this by statements (9.51), (9.52), (9.53), (9.50) and $[K_{\mathbf{u},k} : \mathbb{Q}] \leq D$ we get the estimate

$$h(x_0^{(k)}(\mathbf{u})), h(y_0^{(k)}(\mathbf{u})) \leq \exp\left\{\exp\left\{N^{12}(2d)^{\exp O(r+s)}(h+1)^{8s}\right\}\right\},$$

for every $\mathbf{u} \in \mathcal{S}$ and $k = 1, \ldots, D$, which provides the same upper bound for H_1. Now combining this latter estimate with (9.49) we get the desired bound (9.41). This concludes the proof of Proposition 9.11. □

9.3.4 Concluding the proof of Proposition 9.2

Now Proposition 9.11 also provides upper bounds for the heights of elements of the set $\mathcal{C}_1$. So for any $(x_0, y_0) \in \mathcal{C}_1$ we have

$$\overline{\deg}\, x_0, \overline{\deg}\, y_0 \leq \exp\left\{N^6 (2d)^{\exp O(r+s)} (h+1)^{4s}\right\},$$

and

$$\overline{h}(x_0), \overline{h}(y_0) \leq \exp\left\{\exp\left\{N^{12} (2d)^{\exp O(r+s)} (h+1)^{8s}\right\}\right\},$$

which by Lemma 6.4 concludes the proof of our Proposition 9.2.

Bibliography

[1] M. Aschenbrenner, Ideal membership in polynomial rings over the integers, *J. Amer. Math. Soc.*, **17** (2004), 407–442.

[2] A. Baker, Contributions to the theory of Diophantine equations I. On the representation of integers by binary forms, *Philos. Trans. Roy. Soc. London Ser. A*, **263** (1967/68), 173–191.

[3] A. Baker, Bounds for the solutions of the hyperelliptic equation, *Proc. Cambridge Philos. Soc.*, **65** (1969), 439–444.

[4] A. Baker, *Transcendental number theory*, Cambridge University Press, London-New York, 1975.

[5] A. Baker and G. Wüstholz, *Logarithmic forms and Diophantine geometry*, vol. 9 of *New Mathematical Monographs*, Cambridge University Press, Cambridge, 2007.

[6] A. Bérczes, Effective results for division points on curves in $\mathbb{G}_m^2$, *J. Théor. Nombres Bordeaux*, **27** (2015), 405–437.

[7] A. Bérczes, Effective results for unit points on curves over finitely generated domains, *Math. Proc. Cambridge Phil. Soc.*, **158** (2015), 331–353.

[8] A. Bérczes, J.-H. Evertse and K. Győry, Effective results for linear equations in two unknowns from a multiplicative division group, *Acta Arith.*, **136** (2009), 331–349.

[9] A. Bérczes, J.-H. Evertse and K. Győry, Effective results for hyper- and superelliptic equations over number fields, *Publ. Math. Debrecen*, **82** (2013), 727–756.

[10] A. Bérczes, J.-H. Evertse and K. Győry, Effective results for Diophantine equations over finitely generated domains, *Acta Arith.*, **163** (2014), 71–100.

[11] A. Bérczes, J.-H. Evertse, K. Győry and C. Pontreau, Effective results for points on certain subvarieties of tori, *Math. Proc. Cambridge Phil. Soc.*, **147** (2009), 69–94.

[12] F. Beukers and H. P. Schlickewei, The equation $x + y = 1$ in finitely generated groups, *Acta Arith.*, **78** (1996), 189–199.

[13] F. Beukers and C. J. Smyth, Cyclotomic points on curves, in: *Number theory for the millennium, I (Urbana, IL, 2000)*, A K Peters, Natick, MA, 2002, pp. 67–85.

[14] F. Beukers and D. Zagier, Lower bounds of heights of points on hypersurfaces, *Acta Arith.*, **79** (1997), 103–111.

[15] E. Bombieri, Effective Diophantine approximation on $\mathbf{G}_m$, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, **20** (1993), 61–89.

[16] E. Bombieri and P. B. Cohen, Effective Diophantine approximation on $\mathbb{G}_M$. II, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, **24** (1997), 205–225.

[17] E. Bombieri and P. B. Cohen, An elementary approach to effective Diophantine approximation on $\mathbb{G}_m$, in: *Number theory and algebraic geometry*, Cambridge Univ. Press, Cambridge, 2003, pp. 41–62.

[18] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, Cambridge University Press, Cambridge, 2006.

[19] E. Bombieri and U. Zannier, Algebraic points on subvarieties of $\mathbf{G}_m^n$, *Internat. Math. Res. Notices*, (1995), 333–347.

[20] I. Borosh, M. Flahive, D. Rubin and B. Treybig, A sharp bound for solutions of linear Diophantine equations, *Proc. Amer. Math. Soc.*, **105** (1989), 844–846.

[21] B. Brindza, On S-integral solutions of the equation $y^m = f(x)$, *Acta Math. Hungar.*, **44** (1984), 133–139.

[22] B. Brindza, On the equation $f(x) = y^m$ over finitely generated domains, *Acta Math. Hungar.*, **53** (1989), 377–383.

[23] B. Brindza, The Catalan equation over finitely generated integral domains, *Publ. Math. Debrecen*, **42** (1993), 193–198.

[24] B. Brindza and Á. Pintér, On equal values of binary forms over finitely generated fields, *Publ. Math. Debrecen*, **46** (1995), 339–347.

[25] W. D. Brownawell and D. W. Masser, Vanishing sums in function fields, *Math. Proc. Cambridge Philos. Soc.*, **100** (1986), 427–434.

[26] Y. Bugeaud, Bornes effectives pour les solutions des équations en S-unités et des équations de Thue-Mahler, *J. Number Theory*, **71** (1998), 227–244.

[27] Y. Bugeaud and K. Győry, Bounds for the solutions of unit equations, *Acta Arith.*, **74** (1996), 67–80.

[28] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1959.

[29] J. Coates, An effective p-adic analogue of a theorem of Thue, *Acta Arith.*, **15** (1968/69), 279–305.

[30] J.-H. Evertse, Points on subvarieties of tori, in: *A panorama of number theory or the view from Baker's garden (Zürich, 1999)*, Cambridge Univ. Press, 2002, pp. 214–230.

[31] J.-H. Evertse and K. Győry, Discriminant Equations in Diophantine Number Theory, Cambridge University Press, to appear.

[32] J.-H. Evertse and K. Győry, Effective results for unit equations over finitely generated integral domains, *Math. Proc. Camb. Phil. Soc.*, **154** (2013), 351–380.

[33] J.-H. Evertse and K. Győry, *Unit Equation in Diophantine Number Theory*, Cambridge University Press, 2015.

[34] J.-H. Evertse, H. P. Schlickewei and W. M. Schmidt, Linear equations in variables which lie in a multiplicative group, *Ann. of Math. (2)*, **155** (2002), 807–836.

[35] E. Friedman, Analytic formulas for the regulator of a number field., *Invent. Math.*, **98** (1989), 599–622.

[36] K. Győry, Sur l'irréductibilité d'une classe des polynômes. II, *Publ. Math. Debrecen*, **19** (1972), 293–326.

[37] K. Győry, Sur les polynômes à coefficients entiers et de dicriminant donné II, *Publ. Math. Debrecen*, **21** (1974), 125–144.

[38] K. Győry, On the number of solutions of linear equations in units of an algebraic number field, *Comment. Math. Helv.*, **54** (1979), 583–600.

[39] K. Győry, Bounds for the solutions of norm form, discriminant form and index form equations in finitely generated integral domains, *Acta Math. Hungar.*, **42** (1983), 45–80.

[40] K. Győry, Effective finiteness theorems for polynomials with given discriminant and integral elements with given discriminant over finitely generated domains, *J. Reine Angew. Math.*, **346** (1984), 54–100.

[41] K. Győry and K. Yu, Bounds for the solutions of S-unit equations and decomposable form equations, *Acta Arith.*, **123** (2006), 9–41.

[42] G. Hermann, Die Frage der endlich vielen Schritte in der Theorie der Polynomideale, *Math. Ann.*, **95** (1926), 736–788.

[43] M. Hindry and J. H. Silverman, *Diophantine geometry*, vol. 201 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, 2000.

[44] S. Lang, Integral points on curves, *Inst. Hautes Études Sci. Publ. Math.*, (1960), 27–43.

[45] S. Lang, *Diophantine geometry*, Interscience Tracts in Pure and Applied Mathematics, No. 11, Interscience Publishers (a division of John Wiley & Sons), New York-London, 1962.

[46] S. Lang, Division points on curves, *Ann. Mat. Pura Appl. (4)*, **70** (1965), 229–234.

[47] S. Lang, Report on diophantine approximations, *Bull. Soc. Math. France*, **93** (1965), 177–192.

[48] S. Lang, *Fundamentals of Diophantine geometry*, Springer-Verlag, New York, 1983.

[49] M. Laurent, Equations diophantiennes exponentielles, *Invent. Math.*, **78** (1984), 299–327.

[50] W. LeVeque, On the equation $y^m = f(x)$, *Acta Arith.*, **9** (1964), 209–219.

[51] P. Liardet, Sur une conjecture de Serge Lang, *C. R. Acad. Sci. Paris Sér. A*, **279** (1974), 435–437.

[52] P. Liardet, Sur une conjecture de Serge Lang, in: *Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, Bordeaux, 1974)*, Soc. Math. France, Paris, 1975, pp. 187–210. Astérisque, Nos. 24–25.

[53] S. Louboutin, Explicit bounds for residues of Dedekind zeta functions, values of L-functions at $s = 1$, and relative class numbers, *J. Number Theorey*, **85** (2000), 263–282.

[54] K. Mahler, Zur Approximation algebraischer Zahlen. I, *Math. Ann.*, **107** (1933), 691–730.

[55] R. C. Mason, *Diophantine equations over function fields*, Cambridge University Press, 1984.

[56] E. M. Matveev, An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II (translated from Russian), *Izv. Math.*, **64** (2000), 1217–1269.

[57] C. J. Parry, The $\mathfrak{p}$-adic generalisation of the Thue-Siegel theorem, *Acta Math.*, **83** (1950), 1–100.

[58] C. Pontreau, A Mordell-Lang plus Bogomolov type result for curves in $\mathbb{G}_m^2$, *Monatsh. Math.*, **157** (2009), 267–281.

[59] C. Pontreau, Petits points d'une surface, *Canad. J. Math.*, **61** (2009), 1118–1150.

[60] B. Poonen, Mordell-Lang plus Bogomolov, *Invent. Math.*, **137** (1999), 413–425.

[61] G. Rémond, Décompte dans une conjecture de Lang, *Invent. Math.*, **142** (2000), 513–545.

[62] G. Rémond, Sur les sous-variétés des tores, *Compositio Math.*, **134** (2002), 337–366.

[63] G. Rémond, Approximation diophantienne sur les variétés semi-abéliennes, *Ann. Sci. École Norm. Sup. (4)*, **36** (2003), 191–212.

[64] J. B. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.*, **6** (1962), 64–94.

[65] A. Schinzel and R. Tijdeman, On the equation $y^m = P(x)$, *Acta Arith.*, **31** (1976), 199–204.

[66] H. P. Schlickewei, Lower bounds for heights on finitely generated groups, *Monatsh. Math.*, **123** (1997), 171–178.

[67] W. M. Schmidt, Thue's equation over function fields, *J. Austral. Math. Soc. Ser. A.*, **25** (1978), 385–442.

[68] W. M. Schmidt, *Diophantine approximation*, vol. 785 of *Lecture Notes in Mathematics*, Springer, Berlin, 1980.

[69] W. M. Schmidt, Heights of points on subvarieties of $\mathbf{G}_m^n$, in: *Number theory (Paris, 1993–1994)*, Cambridge Univ. Press, Cambridge, 1996, vol. 235 of *London Math. Soc. Lecture Note Ser.*, pp. 157–187.

[70] C. Siegel, Approximation algebraischer Zahlen, *Math. Z.*, **10** (1921), 173–213.

[71] V. G. Sprindžuk and S. V. Kotov, An effective analysis of the Thue-Mahler equation in relative fields (Russian), *Dokl. Akad. Nauk. BSSR*, **17** (1973), 393–395, 477.

[72] H. M. Stark, Some effective cases of the Brauer-Siegel theorem, *Invent. Math.*, **23** (1974), 135–152.

[73] A. Thue, Über Annäherungswerte algebraischer Zahlen, *J. Reine Angew. Math.*, **135** (1909), 284–305.

[74] L. A. Trelina, S-integral solutions of Diophantine equations of hyperbolic type (in Russian), *Dokl. Akad. Nauk. BSSR*, **22** (1978), 881–884;955.

[75] J. Végső, On superelliptic equations, *Publ. Math. Debrecen*, **44** (1994), 183–187.

[76] P. Voutier, An effective lower bound for the height of algebraic numbers, *Acta Arith.*, **74** (1996), 81–95.

[77] M. Waldschmidt, *Diophantine approximation on linear algebraic groups*, Springer-Verlag, 2000.

[78] K. Yu, P-adic logarithmic forms and group varieties. III, *Forum Math.*, **19** (2007), 187–280.

[79] S. Zhang, Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.*, **8** (1995), 187–221.